

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:14:27 ; Search time 20.83 Seconds

(without alignments)  
1437.746 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598  
Sequence: 1 MAASAMGVLOEREFKKIWT.....KLVRTTEGQWGLFTAKK 494

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_0601:\*

1:	/SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2204	84.8	473	22	AA80627
2	206.5	7.9	285	21	AA70038
3	204.5	7.9	317	20	AAW95016
4	198	7.6	283	20	AA339308
5	193.5	7.4	279	21	AA170041
6	189.5	7.3	275	21	AA32508
7	189.5	7.3	283	21	AA32507
8	188	7.2	293	21	AA321652
9	188	7.2	296	21	AA321526
10	188	7.2	336	21	AA321651
11	188	7.2	339	21	AA321525
					Arabidopsis thalia

12	188	7.2	361	21	AA321524	Arabidopsis thalia
13	188	7.2	366	21	AA321650	Arabidopsis thalia
14	187	7.2	293	21	AA321592	Arabidopsis thalia
15	187	7.2	293	21	AA321595	Arabidopsis thalia
16	187	7.2	336	21	AA321591	Arabidopsis thalia
17	187	7.2	336	21	AA321594	Arabidopsis thalia
18	187	7.2	358	21	AA321595	Arabidopsis thalia
19	187	7.2	365	21	AA321593	Arabidopsis thalia
20	185	7.1	301	22	AA321683	Synechocystis PCC6
21	185	7.1	318	21	AA3216405	Amino acid sequenc
22	185	7.1	318	22	AA3216831	Synechocystis PCC6
23	185	7.1	318	22	AA3216832	Synechocystis PCC6
24	182	7.0	308	21	AA3217857	Arabidopsis thalia
25	182	7.0	344	21	AA3218092	Maize C-24 sterol
26	182	7.0	348	21	AA321856	Arabidopsis thalia
27	179	6.9	280	18	AA3217977	S-adenosyl methion
28	178	6.9	348	20	AA3219017	Arabidopsis gamma-
29	178	6.9	348	21	AA3217937	A. thaliana enviro
30	177	6.8	308	21	AA321382	Arabidopsis thalia
31	177	6.8	348	21	AA321381	Arabidopsis thalia
32	173.5	6.7	266	21	AA321858	Arabidopsis thalia
33	172	6.6	115	21	AA321490	SAM dependent meth
34	170	6.5	352	21	AA321469	Corn gamma-tocophe
35	170	6.5	383	20	AA3210836	Yeast SWT enzyme.
36	169.5	6.5	114	21	AA321486	SAM dependent meth
37	168.5	6.5	266	21	AA321383	Arabidopsis thalia
38	168	6.5	275	20	AA321302	Spnf protein invol
39	165	6.4	349	21	AA3216462	Soybean gamma-toco
40	164.5	6.3	350	21	AA3216472	Soybean gamma-toco
41	162.5	6.3	344	20	AA3210838	Corn SWT enzyme.
42	162.5	6.3	366	21	AA3216473	Soybean gamma-toco
43	162.5	6.3	581	21	AA321804	Arabidopsis thalia
44	160	6.2	361	20	AA3210837	Arabidopsis SWT en
45	159.5	6.1	357	20	AA3214114	Maize cyclopropane

#### ALIGNMENTS

RESULT 1	
AA80627	standard; Protein: 473 AA.
ID	AA80627
AC	AA80627
XX	
DT	02-MAY-2001 (first entry)
XX	
DE	Environmental stress tolerant protein SEQ ID 40.
XX	
KW	Environmental stress resistance; salt; heat; desert; transgenic plant.
XX	
OS	Suada japonica.
XX	
PN	WO200106006-A1.
XX	
PD	25-JAN-2001.
XX	
PE	19-JUL-2000; 2000MO-JP04862.
XX	
PR	19-JUL-1999; 99JP-0235910.
XX	
PR	24-MAR-2000; 2000JP-0085377.
XX	
PA	(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX	
PI	Yamada A, Ozeki Y, Saito T;
XX	
DR	WPI; 2001-147355/15.
XX	
DR	N-PSDB; AAF74206.
XX	
PT	Screening method to obtain DNA encoding environmental stress resistance
XX	
PT	factor, useful for producing transgenic plants resistant to
XX	
XX	environmental stress

PS Claim 64; Page 125-127; 167pp; Japanese.  
 XX  
 CC Polynucleotide sequences AAF74187 - AAF74218 encode proteins  
 CC AAB80608 - AAB80639, which impart environmental stress resistance. The  
 CC invention relates to a method for identifying DNA encoding proteins  
 CC imparting environmental stress resistance. The method comprises inserting  
 CC cDNA from a library originating in a salt-resistant organism into a host  
 CC cell, culturing the transformants under conditions in which the  
 CC untransformed host does not grow well, and selecting for viable clones.  
 CC The method is useful for obtaining DNA encoding environmental stress  
 CC resistance factors. The DNA encoding proteins conferring environmental  
 CC stress resistance, can be used in the production of plants resistant to  
 CC environmental stress, which can be cultivated in unfavourable  
 CC environments such as deserts, salt damaged ground, cold regions and the  
 CC oceans. They can be used for increasing the area of land covered by green  
 CC plants, and desert greening and afforestation. In order to counter the  
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR  
 CC primers AAF74219 and AAF74220 are used in an example illustrating the  
 CC method of the invention.  
 CC  
 XX  
 SO Sequence 473 AA;  
 Query Match 84.8%; Score 2204; DB 22; Length 473;  
 Best Local Similarity 86.4%; Pred. No. 3.9e-198;  
 Matches 408; Conservative 34; Mismatches 30; Indels 0; Gaps 0;  
 22 HSDVITVEAMWLDSDASDLKVEREVLSPYREKSVLELGGIRFGELAKRSQV 81  
 1 hvvdlleammlldsqasdlkveerpellsmprlegkcllelga9lgrfgeleakkg 60  
 82 IALDFIESVIRKKNESINGHYKKNVFCADYSPSLNISPSNVDFIFNNMLMYLSDDEVE 141  
 61 laldlesalalknevinghyknvfcadvtspstsfpsldvifanwlmlylsdeve 120  
 142 RLYERMKWLKPGCYTRFRSCFQSGDHKKSNPTHYRPREFTTKIFKCHMODDSGNS 201  
 121 nlvermlkwlkpg9ylfrfscfngsdhkrknpthyrpreftkafkechldqsgns 180  
 202 YELSLGCKIGAVKSKNONOISMLMOKVDSDDKPFORFLDSQYKFNSTIRYERVF 261  
 181 yelslgckigavvyrnkknqnglswlqkvdskdkqfqlfcltsykcnslltyervf 240  
 262 GPGVSTGLETTEKFEVSKLDLKGQKVLVDCGIGGDFYMAENYVEVVGIDLSIMI 321  
 241 gpgvstggyetctefsmldlkgqkvlvdcgig9gddfyfmaetfvevvgfclsvmi 300  
 322 SPALERSIGLKCAVEFEVADCTKKDIPENSEFDVYSRDTILIHODKPALFRSFHKWLKPG 381  
 301 sfalersiglkcavefadvadctkinypdnsfdvysrdtllhldqkpalfrsfkwlkpg 360  
 382 GKVLISDYCKSAGTSPSAEFAFYIRORGYD-LHDVKAQKMLKDGEVVAENRFDQFIQ 441  
 361 gkvllsdyckkagppspfaaylkrgydlhdvkeygmkldgafvvaenrfdqfllrv 420  
 442 LQKELDALOEKDDFIDFSEEDYNDIVDGWAKALVTRTEGQWGLFIKK 493  
 421 lrnellevkekevlfidfseedyndlvdgwacklrrtaqegwglfvakk 472  
 Db  
 RESULT 2  
 ID AAY70038 standard; protein; 285 AA.  
 XX  
 XX AAY70038;  
 XX  
 XX 05-JUN-2000 (first entry)  
 XX  
 DE A. halophila sarcosine-dimethylglycine methyltransferase.  
 KW Sarcosine-dimethylglycine methyltransferase; SDMT; betaine operon;  
 KM dimethylglycine content; abiotic stress; tolerance; salt; freezing;  
 KW cold; drought stress; pathogenesis-related protein; animal feeding;

KW pH tolerance; environmental stress; inclusion body formation.  
 XX  
 OS Actinopolyspora halophila.  
 XX  
 XX WO200011142-A2.  
 XX  
 XX 02-MAR-2000.  
 XX  
 XX 18-AUG-1999; 99WO-EP06037.  
 XX  
 XX 20-AUG-1998; 98US-0137434.  
 XX  
 XX (CULT-) CULTOR CORP.  
 XX  
 PI Reinikkaenen T, Nyssöelae A, Kerovuo J;  
 DR WPI; 2000-224686/19.  
 DR N-PSDB; AA250960.  
 XX  
 PT New methyltransferases, useful for e.g. producing transgenic plants  
 PT with increased stress tolerance, pathogen resistance or nutritional  
 PT value as animal feed or for improving the viability of microorganisms  
 PT in the food industry  
 XX  
 PS Claim 19; Page 150-152; 176pp; English.  
 XX  
 CC The present sequence is Actinopolyspora halophila sarcosine  
 CC -dimethylglycine methyltransferase (SDMT) catalysing conversion of  
 CC sarcosine to dimethylglycine. Expression vectors comprising the coding  
 CC region from betaine operon can be used to increase intracellular betaine  
 CC and dimethylglycine content. Betaine is accumulated in the cells as a  
 CC result of abiotic stress, and the polynucleotide can be used to increase  
 CC salt, freezing or cold tolerance, increase resistance to drought stress,  
 CC pathogens or induce pathogenesis-related proteins in plants. Transgenic  
 CC organisms can be used as an animal feed ingredient. The polynucleotide  
 CC can also be used to enhance pH tolerance and improve viability of  
 CC organisms when subjected to environmental stress. This can decrease  
 CC inclusion body formation when used in conjunction with polynucleotides  
 CC encoding a heterologous protein.  
 XX  
 SO Sequence 285 AA;  
 Query Match 7.9%; Score 206.5; DB 21; Length 285;  
 Best Local Similarity 26.3%; Pred. No. 4.5e-11;  
 Matches 61; Conservative 44; Mismatches 120; Indels 7; Gaps 4;  
 266 VSTGGLTTEKFEVSKLDLKGQKVLVDCGIGGDFYMAENYVEVVGIDLS--IMISF 323  
 55 latsertvgmagkvvdlsptetrlldlgagyggaarylartgchvclnlseveqnr 114  
 324 ALERSIGLKCAVEFEVADCTKKD--YPENSPDYYSRDTILIHODKPALFRSFHKWLKPG 381  
 115 eltraeglehl--evdgsfedlpygdnafdvwsqdsflhgdstrvneevlrvlkkp 172  
 382 GKVLISDYCKSAGTSPSAEFAFYIRORGYD-LHDVKAQKMLKDGEVVAENRFDQFIQ 440  
 173 gsvlftdpmasdaaknehlplldrlhldslgspgfyfkeltrtlglqnlsefedlseylpv 232  
 441 VLQKELDALOEKDDFIDFSEEDYNDIVDGWAKALVTRTEGQWGLFIKK 492  
 233 hgyrlevlesreneleagfigeeyrahmkcglrnwvgagsgslawgljhar 284  
 Db  
 RESULT 3  
 ID AAW95016 standard; protein; 317 AA.  
 XX  
 XX AAW95016;  
 XX  
 XX 21-MAY-1999 (first entry)  
 XX  
 DE Synechocystis gamma-tocopherol methyltransferase (gamma-TMT).

```

XX  Gamma-tocopherol methyltransferase; gamma-TMT; SLR0089; 165H5T7; meat;
KW  transgenic plant; alpha-tocopherol; gamma-tocopherol; forage plant.
XX
OS  Synecchocystis sp.
XX
PN  WO9904622-A1.
XX
PD  04-FEB-1999.
XX
PF  22-JUL-1998; 98WO-US15137.
XX
PR  17-JUL-1998; 98US-0053819.
PR  25-JUL-1997; 97US-0053819.
PR  26-JAN-1998; 98US-0072497.
XX
PA  (UYNE-) UNIV NEVADA.
XX
PI  Dellapenna D, Shintani DK;
XX
DR  WPI: 1999-142458/12.
DR  N-PSDB: AAX17788.
XX
XX  Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)
PT  methyltransferase coding sequence - useful for producing
PT  ^a-tocopherol, and transgenic plants, seeds and oils with an altered
PT  tocopherol profile
XX
PS  Example 2: Page 33-35; 46pp; English.
XX
XX  The invention provides DNA sequences encoding gamma-tocopherol
CC  methyltransferases (gamma-TMT). The gamma-TMT encoding genes SLR0089
CC  and 165H5T7 are isolated from Synecchocystis and Arabidopsis species
CC  respectively. The DNA fragments are useful for producing transgenic
CC  plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
CC  are also useful for producing alpha-tocopherol and plants with increased
CC  gamma-tocopherol, which may be useful in certain industries such as the
CC  meat industry e.g. for developing forage plants to feed animals. The
CC  production of transgenic plants (and seeds) with a higher alpha-
CC  tocopherol:gamma-tocopherol ratio, using the gamma TMTs, will increase
CC  the level of alpha-tocopherol in the human diet, in addition to enhancing
CC  the stability and shelf life of plants and plant products. Increased
CC  levels of alpha-tocopherol will also increase meat quality and extend
CC  shelf life of post-processed meat products. Plants with a higher alpha-
CC  tocopherol:gamma-tocopherol ratio may also produce advantageous
CC  phenotypes. The present sequence represents a Synecchocystis gamma-TMT.
XX
SO  Sequence 317 AA:

Query Match          7.9%; Score 204.5; DB 20; Length 317;
Best Local Similarity 26.3%; Pred. No. 8.2e-11;
Matches 61; Conservative 48; Mismatches 90; Indels 33; Gaps 8;

OY  215 YVSKKKNQNIISWIMQKVDSEDDKGFORFLDSSQYKENSILKRVFPGYVSTGLET 274
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  35 yeklnfyddssglwedwge-----hmhgyyphngtyidr-----rqqidll 80
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  275 KEF-----VSKDKRGKGVLDVGGIGGDFPMANVDEVWIDLSINMISFALE--RS 328
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  81 kellawapgnasakp-rkildgicgissjylaqqhaewngasispvqeragera 139
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  329 IGLCAVEFEVADCTKKYVENSFVIVSRDILHIOCKPALFRFHHMLKRGKGVLLSD 388
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  140 lglstcgyfyanalldlpfasdsfivwswlesgehmpnkagllgawvllpgrglilat 199
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  389 YC-----KSAGTPEAEFAFYTRQRCYDLH-----DVKAYGMLMDAGFEVY 429
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  200 wchpripdgngpltdaderhll-galidyvclpvyvslpdyaalarecgfge 250
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT  4
AAV39308

```

```

ID  AAV39308 standard; Protein; 283 AA.
XX
AC  AAV39308;
XX
DT  01-DEC-1999 (first entry)
XX
DE  SpnU protein involved in spinosyn biosynthesis.
XX
KW  Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
KW  microtides; arachnid; insect; polyketide; polyketide synthase;
KW  PKS; extender module; initiator module; acyl transferase domain; A1;
KW  acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
KW  dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
KW  insecticide; C-methylation.
XX
OS  Saccharopolyspora spinosa.
XX
PN  WO9946387-A1.
XX
PD  16-SEP-1999.
XX
PF  16-FEB-1999; 99WO-US03212.
XX
PR  09-MAR-1998; 98US-0036987.
XX
PA  (DOMC ) DOW AGROSCIENCES LLC.
XX
PI  Baltz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ;
PI  Treadway PJ, Turner JR, Waldron C;
XX
DR  WPI: 1999-551414/46.
DR  N-PSDB: AAZ21501.
XX
XX  New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
PT  for production of insecticidal spinosyn compounds
XX
XX  Claim 1; Page 156-157; 190pp; English.
XX
XX  This is the amino acid sequence of the product of the spnU gene. The
CC  protein is involved in spinosyn biosynthesis. The SpnU gene is one of 23
CC  genes and open reading frames contained in an 80kb DNA sequence
CC  AA021501. Spinosyns are insecticidal microtides which are useful for the
CC  control of arachnids, nematodes and insects. Biosynthesis of spinosyns
CC  occurs via stepwise condensation and modification of carboxylic acid
CC  precursors generating a linear polyketide which is modified further. The
CC  DNA sequence contains a central region of approximately 55kb which has
CC  homology to the DNA encoding the polyketide synthases (PKS) of known
CC  macrocyclic producers. The spinosyn PKS DNA region consists of 5 ORFs with
CC  stop codons at the end of acyl carrier protein (ACP) domains. Together
CC  the PKS polypeptides (AAV39297-Y39301), form a complex consisting of an
CC  initiator module, spnA, and several extender modules spnB-spnE. The
CC  products of the genes present in the region upstream of the PKS genes
CC  have been assigned names spnF-spnS AAV39302-Y39315 and are responsible
CC  for different modifications in spinosyn biosynthesis. There are also two
CC  ORFs ORF15 and ORF16 present immediately upstream of spnS, producing
CC  polypeptides AAV39316-Y39317, and two ORFs ORF1 and ORF2 present
CC  downstream of the PKS region producing polypeptides AAV39318-Y39319. It
CC  is suggested that SpnU is involved in C-methylation during spinosyn
CC  biosynthesis. The genes are useful to improve yields of spinosyns and
CC  for creating new spinosyns e.g. by mutagenesis, or interruption of steps
CC  in spinosyn biosynthesis. The modified spinosyns may be a new insect
CC  control agent or serve as substrates for further chemical modification
CC  and the creation of new semi-synthetic spinosyns. The genes are also
CC  useful to isolate similar sequences from S. spinosa or other species by
XX  hybridization.
XX
SO  Sequence 283 AA:

Query Match          7.6%; Score 198; DB 20; Length 283;
Best Local Similarity 25.7%; Pred. No. 2.8e-10;
Matches 53; Conservative 44; Mismatches 97; Indels 12; Gaps 5;

```

QY 272 ETTEFEVKLILKPEQGKVLIDVCGIGGSDPEFMANENVDENVGDLSINIMISFA--IERSI 329  
Db :|||::|:|||::|:|||::|:|||::|:|||::|:|||::|:|||::|:|||::|:|||::|:  
38 qitdiftikaalrppahnlfdlsgngqpvrvraacaagvrvgltnvaqnlaaaarlanet 97  
QY 330 GIKCAVEEFVADCTRKDPENSPDVITYSRDITLIHIDOKPALFRSFKMLPGCKVLISDY 389  
Db |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
98 glagslefdldvdgaqlpypdgffgaawamgsvglvdqaairrevhrllepgrfvlgd 157  
QY 390 CKSACTPAAEERAAIYRQGVLDHOVKAKGKLKAGF----VEVIANRN-IDQFIQVL 442  
Db |trvrllpe-eyaaav--wtgltahclnstaalvsasgaelllevtltaqtcrmwswyvdl 214  
443 QKEILDALDEQEKDFIDPFSEEDYNDI 468  
QY ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db 215 lrkldelagvepaavgtyqgyrlgd 240

RESULT 5  
AAAY70041  
ID AAAY70041 standard; protein: 279 AA.  
XX  
XX  
XX AAY70041;  
AC  
XX  
XX  
XX  
DT 05-JUN-2000 (first entry)  
XX  
XX  
De E. halochloris sarcosine-dimethylglycine methyltransferase.  
XX  
KW Sarcosine-dimethylglycine methyltransferase; SDMT; betaine operon;  
KW dimethylglycine content; abiotic stress; tolerance; salt; freezing;  
KW cold; drought stress; pathogenesis-related protein; animal feed;  
XX pH tolerance; environmental stress; inclusion body formation.  
OS Ecotohorhodospira halochloris.  
XX  
PN MO200011142-A2.  
XX  
PD 02-MAR-2000.  
XX  
XX 18-AUG-1999; 99MO-EPO6037.  
XX PF  
PR 20-AUG-1998; 98US-0137434.  
XX  
PA (CULT-) CULTOR CORP.  
PI Reiniikainen T, Nyssoeleae A, Kerovu J;  
DR MPI: 2000-224686/19.  
DR N-PSTDB; AAZ50961.  
DR  
DR  
XX  
PS  
XX

Claim 19; Page 163-164; 176pp; English.

The present sequence is Ecotohorhodospira halochloris sarcosine  
dimethylglycine methyltransferase (SDMT) catalysing conversion of  
sarcosine to dimethylglycine. Expression vectors comprising the coding  
region from betaine operon can be used to increase intracellular betaine  
and dimethylglycine content. Betaine is accumulated in the cells as a  
result of abiotic stress, and the polynucleotide can be used to increase  
salt, freezing or cold tolerance. Increase resistance to drought stress,  
pathogens or induce pathogenesis-related proteins in plants. Transgenic  
organisms can be used as an animal feed ingredient. The polynucleotide  
can also be used to enhance pH tolerance and improve viability of  
organisms when subjected to environmental stress. This can decrease  
inclusion body formation when used in conjunction with polynucleotides  
encoding a heterologous protein.

Sequence 279 AA:

Query Match	7.48;	Score 193.5;	DB 21;	Length 279;
Best Local Similarity	25.48;	Pred. No. 7.3e-10;		
Matches 63;	Conservative 43;	Mismatches 91;	Indels 51;	Gaps 8;

```
OY      274   TREPYSKL--DLKPEQKVLNDGCCGGDGFMAENDEVEVNGIDLSTNMISFALERSIGL    331
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      54   lvermsslsrqlgposyvlldmgagyggsarjlahkcyckvaalhsr--enerdigmnk    111
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY      332   KCQAVE--FEVADCTCKKD--YEENSEFDYVSNDTLIHLIDOKPALFRSFHKMLKPEGKVLIS    387
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      112   egvavhlievdaafedvpyddgyflwsgdsflhspsdreravlreasrvlrsqgeiflc    170
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY      388   D-----CCKSAGCPMSAFNAVINRGVDLHDVKAYGCMKDAEFVE    428
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      172   dpmgaddepegviaplrlhletmtprn-----fyqrllrdlgfee    213
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY      429   VIAENRTDOFTO---VLOKEIDALEOEKDDFIIDSSEEDYNDIVGKNAKKLVRTTESEQO    485
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      214   ltfechtbqlprhygrvreldrregelqghvsaeylemkngjdhw---vngankgylt    270
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY      486   WGLFIATKK    493
          ||| : | : |
Db      271   wgllyfyrk    278
```

RESULT	6
AAB32508	
ID	AAB32508 standard; Protein; 275 AA

DT 19-JAN-2001 (first entry)

DE S. lavendulae Mit N encoded protein sequence.

KM Micomycin; biosynthesis; mitosome ring system; antibiotic; anti-cancer  
 KM anti-inflammatory; immune-enhancer; immunosuppressant; asthma;  
 KM chronic obstructive pulmonary disease; respiratory inflammation;  
 KM fungicide; pesticide.

Streptomyces lavendulae.

PN W0200053737-A2.

PD 14-SEP-2000

PF 10-MAR-2000; 2000WO-US06394

PR 12-MAR-1999; 99US-0266965.

PA (MINU ) UNIV MINNESOTA.

PA (MAOY/) MAO Y.

PA (HEMM/) HE M.

[illegible]

XX  
F1 Shetland DH, MAU 1, VADOGU M, NE M, SHETLAND FC,  
XX

DR WFL; 2000-601380/  
DR N-PSDB: AAC55801.

XX	Novel nucleic acid
PT	

PT cluster useful for cloning mitomycin biosynthetic genes for elucidating the molecular basis of mitomycin system biosynthesis -

Disclosure: Page 354: English

XX  
 XX  
 E

from the mitomycin biosynthetic gene cluster. Mitomycins are a group of

CC amino benzogquinone and axiridine ring systems. The S. lavendulae

CC spanning 55kb of DNA. The invention includes an expression cassette  
 CC comprising a mitomycin biosynthetic gene operably linked to a promoter,  
 CC and host cells transformed with the cassette. The nucleotide, and protein  
 CC sequences and the transformed host cells of the invention result in  
 CC antiasthmatic, antiinflammatory, cytostatic, immunomodulatory, and  
 CC antibiotic activities. The nucleotide sequences are used to elucidate the  
 CC molecular basis for the biosynthesis of the mitosane ring system, as well  
 CC as to engineer the biosynthesis of novel natural products, e.g.  
 CC antibiotics, anti-inflammatory agents, anti-cancer agents,  
 CC immune-enhancers, immunosuppressants, agents to treat asthma, chronic  
 CC obstructive pulmonary disease as well as other disease involving  
 CC respiratory inflammation, or cholesterol-lowering agents or as crop  
 CC protection agents (e.g. fungicides or insecticides) as well as  
 CC biopolymers, e.g., in packaging or biomedical applications, or to engineer  
 CC PHA monomer synthases. Sequences AAC55782-C55881, AAC55815-C55849 and  
 CC AAC32485-B32542 represent mitomycin biosynthetic gene cluster DNA  
 CC sequences and encoded proteins. Sequences AAC55812-C55814,  
 CC AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the  
 CC cloning of the mitomycin biosynthetic genes.

XX Sequence 275 AA;

Query Match 7.3%; Score 189.5; DB 21; Length 275;  
 Best Local Similarity 27.5%; Pred. No. 1.7e-09;  
 Matches 61; Conservative 38; Mismatches 102; Indels 21; Gaps 7;

OY 263 PGYSTGLETTEKFEVSLDKPGOKVLDVCGGDFYMAENYDEVVGIDLSINNIS 322  
 DB 41 pttveeaadrltdyvsrlrlapgervidvgsgngkatrlaarlgratary--slpqy 98  
 OY 323 FALERSTLAK--CAVEFEVADCTKKDYPENSFVYISRDITLHIDKPALEFRSFHKLK 379  
 DB 99 vglstrqlaekgdeeterrigdmalprpdsidacysaleschaleradrvftelarlr 158  
 OY 380 PGKVLISDY-----CKSAGTPSAEFAAYIRORYDLHDVAKYGMKLDAG-----EVEV 429  
 DB 159 pggrvtvtdlvtlrpldsarltvtdandnfqgspvl--treayedcmrsvglevefidi 217  
 OY 430 IAEKNTDOFIOVLOKELDALBEKEDFDIDFSEEDYNIVDVG 471  
 DB 218 gdevrps--yeavaakmraardelsglsmdd---eafhmrvdg 254

RESULT 7

AAB32507 ID AAB32507 standard; Protein: 283 AA.

AC AAB32507;

DT 19-JAN-2001 (first entry)

XX S. lavendulae Mit M encoded protein sequence.

XX Mitomycin; biosynthesis; mitosane ring system; antibiotic; anti-cancer;

XX anti-inflammatory; immune-enhancer; immunosuppressant; asthma;

XX chronic obstructive pulmonary disease; respiratory inflammation;

XX fungicide; pesticide.

OS Streptomyces lavendulae.

XX WO200053737-A2.

XX 14-SEP-2000.

XX 10-MAR-2000; 2000WO-US06394.

XX 12-MAR-1999; 99US-0266965.

XX (MINU ) UNIV MINNESOTA.

XX (SHER/) SHERMAN D H.

XX (MAOY/) MAO Y.

XX (VARO/) VAROGLU M.

PA (HEM/) HE M.  
 PA (SHEL/) SHELDON P C.  
 XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;  
 DR WPI: 2000-601980/57.  
 XX N-PSDB: AAC55802.

PT Novel nucleic acid molecule comprising mitomycin biosynthetic gene  
 PT cluster useful for cloning mitomycin biosynthetic genes for elucidating  
 PT the molecular basis of mitosane ring system biosynthesis

PS Disclosure: Page 353; 399pp; English.

CC This invention relates to isolated and purified nucleic acid molecules  
 CC from the mitomycin biosynthetic gene cluster. Mitomycins are a group of  
 CC natural products that contain a variety of functional groups, including  
 CC amino benzoxazinone and axiridine ring systems. The S. lavendulae  
 CC mitomycin biosynthetic gene cluster comprises 47 mitomycin genes  
 CC spanning 55kb of DNA. The invention includes an expression cassette  
 CC comprising a mitomycin biosynthetic gene operably linked to a promoter,  
 CC and host cells transformed with the cassette. The nucleotide, and protein  
 CC sequences and the transformed host cells of the invention result in  
 CC antiasthmatic, antiinflammatory, cytostatic, immunomodulatory, and  
 CC antibiotic activities. The nucleotide sequences are used to elucidate the  
 CC molecular basis for the biosynthesis of the mitosane ring system, as well  
 CC as to engineer the biosynthesis of novel natural products, e.g.  
 CC antibiotics, anti-inflammatory agents, anti-cancer agents,  
 CC immune-enhancers, immunosuppressants, agents to treat asthma, chronic  
 CC obstructive pulmonary disease as well as other disease involving  
 CC respiratory inflammation, or cholesterol-lowering agents or as crop  
 CC protection agents (e.g. fungicides or insecticides) as well as  
 CC biopolymers, e.g., in packaging or biomedical applications, or to engineer  
 CC PHA monomer synthases. Sequences AAC55782-C55881, AAC55815-C55849 and  
 CC AAC32485-B32542 represent mitomycin biosynthetic gene cluster DNA  
 CC sequences and encoded proteins. Sequences AAC55812-C55814,  
 CC AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the  
 CC cloning of the mitomycin biosynthetic genes.

XX Sequence 283 AA;

Query Match 7.3%; Score 189.5; DB 21; Length 283;  
 Best Local Similarity 25.4%; Pred. No. 1.8e-09;  
 Matches 60; Conservative 52; Mismatches 93; Indels 31; Gaps 9;

OY 274 TKEFYSKLDIKRGOKVLDVCGGDFYMAENYDEVVGIDLSINMI--SFALERSTGL 331  
 DB 60 tdmmaeerlrigasrvldigcgvqtpgyvrlasgahvtgjsvshqvyrranaleeaql 119  
 OY 332 KCAVEFEVADCTKKDYPENSFVYISRDITLHIDKPALEFRSFHKLKPGKVLISDYCK 391  
 DB 120 adrarfrgradamdlpdefdsifavlalesilmprragvlagvgvrlppgglvltldffe 179  
 OY 332 SAGTPSAEFAAYIRORYDLHD-----VKAYGMKLDAG-----EVEVIAENRTDQFI 439  
 DB 180 raplapgeaavqy----lhdftmtbmsaeayppllrigagjwleefld-lsdqtlektf 234  
 OY 440 QVLOKELDALBEKEDFDIDFSEEDYNIVDGMKAKIVRTTGEQOMG--LEFIARK 493  
 DB 235 rliserlinskkrlle---tqigeemvngfdpg---dlv---gvkexfylllvagr 280

RESULT 8

AAG21652 ID AAG21652 standard; Protein: 293 AA.

AC AAG21652;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 24284.

Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

Arabidopsis thaliana.

EPI033405-A2.

PN  
XX  
XX  
PD  
06-SEP-2000.

XX  
PF  
25-FEB-2000; 2000EP-0301439.

XX  
PR  
25-FEB-1999; 990S-0121825.

PR  
05-MAR-1999; 990S-0123380.

PR  
09-MAR-1999; 990S-0123548.

PR  
23-MAR-1999; 990S-0125788.

PR  
25-MAR-1999; 990S-0126284.

PR  
29-MAR-1999; 990S-0126785.

PR  
01-APR-1999; 990S-0127462.

PR  
06-APR-1999; 990S-0128234.

PR  
08-APR-1999; 990S-0128714.

PR  
16-APR-1999; 990S-0129845.

PR  
19-APR-1999; 990S-0130077.

PR  
21-APR-1999; 990S-0130449.

PR  
23-APR-1999; 990S-0130510.

PR  
28-APR-1999; 990S-0130891.

PR  
30-APR-1999; 990S-0131449.

PR  
30-APR-1999; 990S-0132048.

PR  
04-MAY-1999; 990S-0132407.

PR  
04-MAY-1999; 990S-0132484.

PR  
05-MAY-1999; 990S-0132485.

PR  
06-MAY-1999; 990S-0132486.

PR  
07-MAY-1999; 990S-0132487.

PR  
11-MAY-1999; 990S-0132863.

PR  
14-MAY-1999; 990S-0134218.

PR  
14-MAY-1999; 990S-0134219.

PR  
14-MAY-1999; 990S-0134221.

PR  
14-MAY-1999; 990S-0134370.

PR  
18-MAY-1999; 990S-0134768.

PR  
19-MAY-1999; 990S-0134841.

PR  
20-MAY-1999; 990S-0135324.

PR  
21-MAY-1999; 990S-0135353.

PR  
24-MAY-1999; 990S-0135629.

PR  
25-MAY-1999; 990S-0136021.

PR  
27-MAY-1999; 990S-0136392.

PR  
28-MAY-1999; 990S-0136782.

PR  
01-JUN-1999; 990S-0137222.

PR  
03-JUN-1999; 990S-0137528.

PR  
04-JUN-1999; 990S-0137502.

PR  
07-JUN-1999; 990S-0137724.

PR  
08-JUN-1999; 990S-0138094.

PR  
10-JUN-1999; 990S-0138540.

PR  
10-JUN-1999; 990S-0138847.

PR  
14-JUN-1999; 990S-0139119.

PR  
16-JUN-1999; 990S-0139452.

PR  
16-JUN-1999; 990S-0139453.

PR  
17-JUN-1999; 990S-0139492.

PR  
18-JUN-1999; 990S-0139454.

PR  
18-JUN-1999; 990S-0139455.

PR  
18-JUN-1999; 990S-0139456.

PR  
18-JUN-1999; 990S-0139457.

PR  
18-JUN-1999; 990S-0139458.

PR  
18-JUN-1999; 990S-0139459.

PR  
18-JUN-1999; 990S-0139460.

PR  
18-JUN-1999; 990S-0139461.

PR  
18-JUN-1999; 990S-0139462.

PR  
18-JUN-1999; 990S-0139463.

PR  
18-JUN-1999; 990S-0139750.

PR  
18-JUN-1999; 990S-0139763.

PR  
21-JUN-1999; 990S-0139817.

PR  
22-JUN-1999; 990S-0139899.

PR  
23-JUN-1999; 990S-0140353.

PR  
23-JUN-1999; 990S-0140354.

PR  
24-JUN-1999; 990S-0140695.

PR  
28-JUN-1999; 990S-0140823.

PR  
29-JUN-1999; 990S-0140991.

PR  
30-JUN-1999; 990S-0141287.

PR  
01-JUL-1999; 990S-0141842.

PR  
02-JUL-1999; 990S-0142154.

PR  
02-JUL-1999; 990S-0142055.

PR  
06-JUL-1999; 990S-0142390.

PR  
08-JUL-1999; 990S-0142803.

PR  
09-JUL-1999; 990S-0142920.

PR  
12-JUL-1999; 990S-0142977.

PR  
13-JUL-1999; 990S-0143542.

PR  
14-JUL-1999; 990S-0143624.

PR  
15-JUL-1999; 990S-0144005.

PR  
16-JUL-1999; 990S-0144005.

PR  
16-JUL-1999; 990S-0144085.

PR  
16-JUL-1999; 990S-0144086.

PR  
19-JUL-1999; 990S-0144325.

PR  
19-JUL-1999; 990S-0144331.

PR  
19-JUL-1999; 990S-0144332.

PR  
19-JUL-1999; 990S-0144333.

PR  
19-JUL-1999; 990S-0144334.

PR  
19-JUL-1999; 990S-0144335.

PR  
20-JUL-1999; 990S-0144352.

PR  
20-JUL-1999; 990S-0144632.

PR  
20-JUL-1999; 990S-0144884.

PR  
21-JUL-1999; 990S-0144814.

PR  
21-JUL-1999; 990S-0145086.

PR  
21-JUL-1999; 990S-0145088.

PR  
22-JUL-1999; 990S-0145085.

PR  
22-JUL-1999; 990S-0145087.

PR  
22-JUL-1999; 990S-0145089.

PR  
22-JUL-1999; 990S-0145192.

PR  
23-JUL-1999; 990S-0145145.

PR  
23-JUL-1999; 990S-0145218.

PR  
23-JUL-1999; 990S-0145224.

PR  
26-JUL-1999; 990S-0145226.

PR  
27-JUL-1999; 990S-0145913.

PR  
27-JUL-1999; 990S-0145918.

PR  
27-JUL-1999; 990S-0145919.

PR  
28-JUL-1999; 990S-0145951.

PR  
02-AUG-1999; 990S-0146386.

PR  
02-AUG-1999; 990S-0146388.

PR  
02-AUG-1999; 990S-0146389.

PR  
03-AUG-1999; 990S-0147038.

PR  
04-AUG-1999; 990S-0147204.

PR  
04-AUG-1999; 990S-0147302.

PR  
05-AUG-1999; 990S-0147192.

PR  
05-AUG-1999; 990S-0147260.

PR  
06-AUG-1999; 990S-0147303.

PR  
06-AUG-1999; 990S-0147416.

PR  
09-AUG-1999; 990S-0147493.

PR  
09-AUG-1999; 990S-0147935.

PR  
10-AUG-1999; 990S-0148171.

PR  
11-AUG-1999; 990S-0148319.

PR  
12-AUG-1999; 990S-0148341.

PR  
13-AUG-1999; 990S-0148565.

PR  
13-AUG-1999; 990S-0148684.

PR  
16-AUG-1999; 990S-0149368.

PR  
17-AUG-1999; 990S-0149175.

PR  
18-AUG-1999; 990S-0149426.

PR  
20-AUG-1999; 990S-0149722.

PR  
20-AUG-1999; 990S-0149723.

PR  
20-AUG-1999; 990S-0149929.

PR  
23-AUG-1999; 990S-0149802.

PR  
23-AUG-1999; 990S-0149930.

PR  
25-AUG-1999; 990S-0150566.

PR  
26-AUG-1999; 990S-0150884.

PR  
27-AUG-1999; 990S-0151065.

PR  
27-AUG-1999; 990S-0151066.

PR  
27-AUG-1999; 990S-0151080.

PR  
30-AUG-1999; 990S-0151303.

PR  
31-AUG-1999; 990S-0151438.

PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156536.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 7.28; Score 188; DB 21; Length 293;  
Best Local Similarity 27.38; Pred. No. 2.6e-09;

Matches 65; Conservative 36; Mismatches 93; Indels 44; Gaps 7;

OY 280 KLIDKPGKVDVDCGCGSGDFYNAENVYVNCIDLSINNISFA--LERSTGLKCAVEF 337  
DB 46 q1g1qpgqkvdvdcg1qgplre1arf1nsavtqlnnneyq1ltqkclnrlagvdkcnf 105  
OY 338 EVADCTKDYDENSEFDVYSRDTILHIDQKPALEFRSPFKWLKPGCKVLISDYCKS-AGTP 396  
DB 106 vkadftmmpfensfdvayaleatchapdaygcylrylv1kpggcfaayekwmtdafcp 165  
OY 397 SAEFAAYLR--QNGVDLHDVKAYGK--MLKDAGFEVIAENR----- 434  
DB 166 dnaehqk1gk1g1dgydp1r1ltk1c1eal1kgagf-ev1ewek1akd1aspvpy1pldkn 224  
OY 435 -----TQGFIOVLQKELDLEQEKDPIDFSESDYVDYDGMKAKL 476  
DB 225 hfs1ssfltlavgrf1tknmvkl1--ey1rlapqsg1rsvn1flqeaeg1vdggr1rel 280

RESULT 9  
AAG21526  
ID AAG21526 standard; Protein: 296 AA.  
XX  
AC AAG21526;

XX XX  
DT 17-OCT-2000 (first entry)  
XX XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24113.  
XX XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PE 25-FEB-2000; 2000EP-0301439.  
XX XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 19-APR-1999; 99US-0129845.  
PR 21-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.

```
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145142.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147203.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
```

```
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161999.
PR 28-OCT-1999; 99US-0162142.
```

Query Match 7.2%; Score 188; DB 21; Length 296;  
Best Local Similarity 27.3%; Pred. No. 2.6e-09;  
Matches 65; Conservative 36; Mismatches 93; Indels 44; Gaps 7;

```
QY 280 KLDKPGOKVLVDGCGIGGIDFYAENYDVEVVGIDISIMNISPA--LEBSIGLKAVEF 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 46 qldgpgqkvlvdgvggldrelarfsnsavqldmneqdlrtgkelnlagvdkcnf 105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 338 EVADCTKKDYPENSDVYISRDITLHIQDKPALFRSHKMKLPGKGYLIDYCKS-AGTP 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 vkadfmkmpfensfdavyleatchapdaygcykelyrvlkpgqcfayewcmtdfdp 165
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 397 SAEEFAAYIR---QRGYLVHGVKAVGK--MLKDAGFYEVTAENR----- 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 166 dnaehqkikgeielgdglpirltltkclealkqagf-evlwekdlaakdspvpywlp]dkn 224
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 435 -----TDQFIQVLOKELDALOEKDDFIDDFSEEDYNDIVDGMKAKL 476
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 hfslsfrltavgrftlkmwkl--eyrlapqsgsrvenfliegaaeglvddggrrei 280
```



RESULT 10  
AC21651  
ID AAG21651 standard; Protein; 336 AA.  
XX  
AC AAG21651;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24283.  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.

PR	20-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149920.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154772.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157863.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

	Query Match	7.2%	Score 188;	DB 21;	Length 336;	
	Best Local Similarity	27.3%*	Pred. No. 3.2e-09;			
	Matches 65; Conservative 36;	Mismatches 93;	Indels 44;	Gaps 7;		
OY	280 KLDELPGKGVLDVGGGIGGDFPEMAENVDVEVGIDLSTINMSFA--LERSIGLKCAVEF	337				
	: : : : :   : : : : :   : : : : :					
DQ	89 qqlgldpgskvldvgcgslggplreiairfnsavtglmneeyqdltrgelnlrllagvaktcnf	148				
OY	338 EVADCTKRDPENSPFDVIYSRDTILHIDKPALFRSFHKMLKPGSKVLISDYCKS-AGRP	396				
	: : : : :   : : : : :   : : : : :					
DQ	149 vkadfmkmpfnspfsfdavaylaieatchapdagycyleiyvlrkpgcfaayewcmtdafdp	208				
OY	397 SAEPAYIR---QRGYDLHDVKATYK---MLKDAGFEVYIAENR-----	434				
	: : : : :   : : : : :   : : : : :					
DQ	209 dnaehqkkkgelsgldgldpdrlrttkckjealkqaef-evliewekdlaakspvwpylpdkn	267				

QY	435	-----TPOGFOVQLKEKELDALEOEKDEIDFSESEDVNDIVDGWKATL	476
Db	268	hfsjssffltavgrfllkmwkl--eyirlapgsqgrvnflqgaagjvdygrrei	323
RESULT 11			
ID	AGG21525		
AC	AGG21525 standard; Protein; 339 AA.		
XX	AGG21525;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 24112.		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126284.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	16-APR-1999; 99US-0128714.		
PR	19-APR-1999; 99US-0129845.		
PR	21-APR-1999; 99US-0130077.		
PR	23-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	23-APR-1999; 99US-0130891.		
PR	28-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	30-APR-1999; 99US-0132407.		
PR	04-MAY-1999; 99US-0132484.		
PR	05-MAY-1999; 99US-0132485.		
PR	06-MAY-1999; 99US-0132486.		
PR	06-MAY-1999; 99US-0132487.		
PR	07-MAY-1999; 99US-0132863.		
PR	11-MAY-1999; 99US-0134256.		
PR	14-MAY-1999; 99US-0134218.		
PR	14-MAY-1999; 99US-0134219.		
PR	14-MAY-1999; 99US-0134221.		
PR	14-MAY-1999; 99US-0134370.		
PR	18-MAY-1999; 99US-0134768.		
PR	19-MAY-1999; 99US-0134941.		
PR	20-MAY-1999; 99US-0135124.		
PR	21-MAY-1999; 99US-0135353.		
PR	24-MAY-1999; 99US-0135629.		
PR	25-MAY-1999; 99US-0136021.		
PR	27-MAY-1999; 99US-0136392.		
PR	28-MAY-1999; 99US-0136782.		
PR	01-JUN-1999; 99US-0137222.		
PR	03-JUN-1999; 99US-0137528.		
PR	04-JUN-1999; 99US-0137502.		
PR	07-JUN-1999; 99US-0137724.		
PR	08-JUN-1999; 99US-0138094.		
PR	10-JUN-1999; 99US-0138540.		
PR	10-JUN-1999; 99US-0138847.		
PR	14-JUN-1999; 99US-0139119.		
PR	16-JUN-1999; 99US-0139452.		
PR	17-JUN-1999; 99US-0139453.		
PR	17-JUN-1999; 99US-0139452.		

```
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151920.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 7.2%; Score 188; DB 21; Length 339;
Best Local Similarity 27.3%; Pred. No. 3.2e-09;
Matches 65; Conservative 36; Mismatches 93; Indels 44; Gaps 7;
```



PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	25-AUG-1999;	99US-0149930.
PR	26-AUG-1999;	99US-0150566.
PR	27-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 7.2%; Score 188; DB 21; Length 361;  
Best Local Similarity 27.3%; Pred. No. 3.5e-09;  
Matches 65; Conservative 36; Indels 44; Gaps 7,  
  
QY 280 KIDLPGKVLDDVGGIGGDFEYMENDEYVAGDLSINMISFA--LEERSIGLKCAVEF 337  
|:::||||||| |::|::|::|::|::|

```
Db      111 qtlqbpagkvldvgvsglbgplrelarfensavtglmnegyiltgkclnrlagvdkcnf 170
Qy      338 EVADCTKDDVPENSEVDVYISRDFTIIHQDPALFRSPFAKMLPGCGKVLIDSYCKS-AGTP 396
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      171 vkadfmknmpfensfdavdayeatchapdaygcykelyrvlkpgcfcaayewcmtdafdp 230
Qy      397 SAEFNAAYR---QRGVLDHDKVAKGCK---MKDAGFEVETLENR-----434
Db      231 dnaehqklkgeletgdglpdlrllrtckclealkgaqf-evlewelakdsprwylpdln 289
Qy      435 -----TQGFIVLQKELEDALEOEKKDFIDPSESDYNDIYDGWKAKL 476 -
Db      290 hfslsfirltavgrfltknmwkiil--eylrIapgsgsrvsnfleqaagevlvdggrrerl 345

RESULT   13
AAG21650
ID       AAG21650 standard; Protein: 366 AA.
XX
AC       AAG21650;
XX
DT       17-OCT-2000 (first entry)
XX
DE       Arabidopsis thaliana protein fragment SEQ ID NO: 24282.
XX
KW       Protein identification; signal transduction pathway; metabolic pathway;
KM       hybridisation assay; genetic mapping; gene expression control; promoter;
KN       termination sequence.
XX
OS       Arabidopsis thaliana.
XX
PN       EP1033405-A2.
XX
PD       06-SEP-2000.
XX
PE       25-FEB-2000; 2000EP-0301439.
XX
PR       25-FEB-1999; 99US-0121825.
XX       05-MAR-1999; 99US-0123180.
PR       09-MAR-1999; 99US-0123548.
PR       23-MAR-1999; 99US-0125788.
PR       25-MAR-1999; 99US-0126264.
PR       29-MAR-1999; 99US-0126785.
PR       01-APR-1999; 99US-0127462.
PR       06-APR-1999; 99US-0128234.
PR       08-APR-1999; 99US-0128714.
PR       16-APR-1999; 99US-0129845.
PR       19-APR-1999; 99US-0130077.
PR       21-APR-1999; 99US-0130449.
PR       23-APR-1999; 99US-0130510.
PR       23-APR-1999; 99US-0130891.
PR       28-APR-1999; 99US-0131449.
PR       30-APR-1999; 99US-0132048.
PR       30-APR-1999; 99US-0132407.
PR       04-MAY-1999; 99US-0132484.
PR       05-MAY-1999; 99US-0132485.
PR       06-MAY-1999; 99US-0132486.
PR       06-MAY-1999; 99US-0132487.
PR       07-MAY-1999; 99US-0132863.
PR       11-MAY-1999; 99US-0134256.
PR       14-MAY-1999; 99US-0134218.
PR       14-MAY-1999; 99US-0134219.
PR       14-MAY-1999; 99US-0134221.
PR       14-MAY-1999; 99US-0134370.
PR       14-MAY-1999; 99US-0134376.
PR       18-MAY-1999; 99US-0134941.
PR       19-MAY-1999; 99US-0135124.
PR       20-MAY-1999; 99US-0135353.
PR       21-MAY-1999; 99US-0135629.
PR       24-MAY-1999; 99US-0136021.
PR       25-MAY-1999; 99US-0136392.
PR       27-MAY-1999; 99US-0136782.
PR       28-MAY-1999; 99US-0137222.
```

PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144684.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.

PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148664.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149475.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 11-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154479.  
PR 20-SEP-1999; 9905-0154739.  
PR 22-SEP-1999; 9905-0155139.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158399.  
PR 12-OCT-1999; 9905-0158493.  
PR 13-OCT-1999; 9905-0159293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159295.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159332.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.  
PR 21-OCT-1999; 9905-0160768.  
PR 21-OCT-1999; 9905-0160770.  
PR 21-OCT-1999; 9905-0160814.  
PR 21-OCT-1999; 9905-0160815.  
PR 22-OCT-1999; 9905-0160880.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160989.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161406.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.  
PR 28-OCT-1999; 9905-0161992.  
PR 28-OCT-1999; 9905-0161993.  
PR 29-OCT-1999; 9905-0162142.

Query Match

7.2%; Score 188; DB 21; Length 366;

	Best Local Similarity	27.3%.	Pred. NO. 3.6e-09;	
	Matches	65;	Conservative	36; Mismatches 93; Indels 44; Gaps 7;
Oy	280	KLDLPGOKVLVDVGGCGDGYMAENYDEVVEVIGDLISNMISFA--LERSTGLKCAVEF	337	
	:	: : : : : : : : : : : :	:	:
Dd	119	qIgiPgqkvlvdvgsIggIrrelarIfmsavtgtInnneyqlrtgkrlnarlgvdcItcf	178	
Oy	338	EVAADTKKDIDPNSDVIVSRDTLIHIDDKPALPFSPFHMKLRPGGVNLSDYCKS-ACNP	396	
	:	: : : : : : : : : : : :	:	:
Dd	179	vktadtmkpmpfpenstcdavayaleatcthapdaygcyeIlyvlrpggcfaeyecmIdafcp	238	
Oy	397	SAEPAAYIR---QRGTDLHDVKAYEK--MLKDAGFEVIYAENR-----	434	
	:	: : : : : : : : : : : :	:	:
Dd	239	dnaehqklkgeleIgdlpdrIlttckleajlkqaif-evitwekdIakdsppvpwylpldkn	297	
Oy	435	-----TDOFIQVLQKEIDALBQEKKDFIDDPSEEDYNIVGWAKTL	476	
	:	: : : : : : : : : : : :	:	:
Dd	298	hfslasffrltavgfrfktnmwkiI--evlrlapqsgsqvrnsflegaaelvgdqreI	353	

RESULT	14
AAG45952	
ID	AAG45952 standard; Protein; 293 AA.

AC AAG45952;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 57754.

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

OS *Arabidopsis thaliana*.

PN EP1033405-A2

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

25-FEB-1999;	9905-012182.5.
PR	
05-MAR-1999;	9905-012318.0.
PR	
09-MAR-1999;	9905-012354.8.
PR	
23-MAR-1999;	9905-012578.8.
PR	
25-MAR-1999;	9905-012626.4.
PR	
29-MAR-1999;	9905-012678.5.
PR	
01-APR-1999;	9905-012746.2.
PR	
06-APR-1999;	9905-012823.4.
PR	
08-APR-1999;	9905-012871.4.
PR	
16-APR-1999;	9905-012984.5.
PR	
19-APR-1999;	9905-013007.7.
PR	
21-APR-1999;	9905-013044.9.
PR	
23-APR-1999;	9905-013051.0.
PR	
23-APR-1999;	9905-013089.1.
PR	
28-APR-1999;	9905-013144.9.
PR	
30-APR-1999;	9905-013204.8.
PR	
30-APR-1999;	9905-013240.7.
PR	
04-MAY-1999;	9905-013248.4.
PR	
05-MAY-1999;	9905-013248.5.
PR	
06-MAY-1999;	9905-013248.6.
PR	
06-MAY-1999;	9905-013248.7.
PR	
07-MAY-1999;	9905-013286.3.
PR	
11-MAY-1999;	9905-013425.6.
PR	
14-MAY-1999;	9905-013421.8.
PR	
14-MAY-1999;	9905-013421.9.
PR	
14-MAY-1999;	9905-013422.1.
PR	
14-MAY-1999;	9905-013437.0.
PR	
18-MAY-1999;	9905-013476.8.
PR	
19-MAY-1999;	9905-013494.1.
PR	
20-MAY-1999;	9905-013512.4.
PR	
21-MAY-1999;	9905-013535.3.

PR	24-MAY-1999	9905-01356529
PR	25-MAY-1999	9905-01366021
PR	27-MAY-1999	9905-01363692
PR	28-MAY-1999	9905-01367782
PR	01-JUN-1999	9905-01376722
PR	03-JUN-1999	9905-01375728
PR	04-JUN-1999	9905-01375702
PR	07-JUN-1999	9905-01377124
PR	08-JUN-1999	9905-01380094
PR	10-JUN-1999	9905-01385400
PR	10-JUN-1999	9905-01388147
PR	14-JUN-1999	9905-01391819
PR	16-JUN-1999	9905-01394512
PR	16-JUN-1999	9905-01394533
PR	17-JUN-1999	9905-01394692
PR	18-JUN-1999	9905-01394612
PR	18-JUN-1999	9905-01394635
PR	18-JUN-1999	9905-01397503
PR	18-JUN-1999	9905-01397630
PR	21-JUN-1999	9905-01398177
PR	22-JUN-1999	9905-01398699
PR	23-JUN-1999	9905-01403553
PR	23-JUN-1999	9905-01406354
PR	24-JUN-1999	9905-01405095
PR	28-JUN-1999	9905-01408923
PR	29-JUN-1999	9905-01409821
PR	30-JUN-1999	9905-01412877
PR	01-JUL-1999	9905-01412854
PR	01-JUL-1999	9905-01412842
PR	02-JUL-1999	9905-01420055
PR	06-JUL-1999	9905-01423900
PR	08-JUL-1999	9905-01428203
PR	09-JUL-1999	9905-01428220
PR	12-JUL-1999	9905-01429277
PR	13-JUL-1999	9905-01435324
PR	14-JUL-1999	9905-01435624
PR	15-JUL-1999	9905-01440005
PR	16-JUL-1999	9905-01440085
PR	16-JUL-1999	9905-01440086
PR	19-JUL-1999	9905-01443425
PR	19-JUL-1999	9905-01443433
PR	19-JUL-1999	9905-01443432
PR	19-JUL-1999	9905-01443433
PR	19-JUL-1999	9905-01443432
PR	20-JUL-1999	9905-01443432
PR	20-JUL-1999	9905-01443432
PR	21-JUL-1999	9905-01445086
PR	21-JUL-1999	9905-01450088
PR	22-JUL-1999	9905-01450087
PR	22-JUL-1999	9905-01450089
PR	22-JUL-1999	9905-01451592
PR	23-JUL-1999	9905-01451545
PR	23-JUL-1999	9905-01455224
PR	26-JUL-1999	9905-01455276
PR	27-JUL-1999	9905-01455313
PR	27-JUL-1999	9905-01455118
PR	28-JUL-1999	9905-01459511
PR	02-AUG-1999	9905-01465388





PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135621.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136332.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138004.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140655.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144375.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 05-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:16:42 ; Search time 13.12 Seconds  
(Without alignments)  
775.277 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598

Sequence: 1 MAASAMGVLOEREVFKKYYMI.....KIVRTGEQOGLFIARKM 494

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA\*  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	207	8.0	363	4	US-09-041-718-5
2	198	7.6	283	4	US-09-036-987A-13
3	179	6.9	280	4	US-09-029-603-2
4	172.5	6.6	282	1	US-08-457-245-9
5	170	6.5	383	4	US-09-041-718-3
6	168	6.5	275	4	US-09-036-987A-7
7	160	6.2	361	4	US-09-041-718-4
8	148.5	5.7	376	4	US-09-041-718-2
9	128.5	4.9	276	3	US-08-935-263-16
10	128.5	4.9	308	1	US-08-457-245-8
11	122.5	4.7	1086	6	5386025-8
12	120	4.6	1084	1	US-08-455-543A-56
13	120	4.6	1084	2	US-08-223-305C-56
14	120	4.6	1103	1	US-08-455-543A-53
15	120	4.6	1103	2	US-08-223-305C-53
16	115	4.4	287	1	US-08-457-245-3
17	114.5	4.4	1091	1	US-07-745-206A-25
18	114.5	4.4	1091	1	US-08-455-543A-52
19	114.5	4.4	1091	2	US-08-223-305C-52
20	114.5	4.4	1091	2	US-08-311-363-25
21	114.5	4.4	1091	4	US-08-713-118-4
22	114.5	4.4	1091	4	US-09-452-007-4
23	114	4.4	1079	1	US-08-455-543A-55
24	114	4.4	1079	2	US-08-223-305C-55
25	113.5	4.4	1106	1	US-08-435-675B-5
26	110.5	4.3	1106	1	US-08-336-257A-8
27	108.5	4.2	1086	1	US-08-455-543A-54

28	108.5	4.2	1086	2	US-08-223-305C-54	Sequence 54, App1
29	105	4.0	715	2	US-08-484-993B-10	Sequence 10, App1
30	105	4.0	715	2	US-08-484-158B-10	Sequence 10, App1
31	105	4.0	715	2	US-08-484-596A-10	Sequence 10, App1
32	105	4.0	715	2	US-08-480-150A-10	Sequence 10, App1
33	105	4.0	715	3	US-08-458-731-10	Sequence 10, App1
34	105	4.0	715	3	US-08-149-223A-10	Sequence 10, App1
35	104.5	4.0	715	3	US-08-961-083-64	Sequence 64, App1
36	102	3.9	773	2	US-08-966-389-4	Sequence 4, App1
37	102	3.9	773	2	US-09-103-509-4	Sequence 4, App1
38	102	3.9	773	2	US-09-102-644-4	Sequence 4, App1
39	102	3.9	773	2	US-09-218-032-4	Sequence 4, App1
40	100	3.8	460	3	US-09-147-522-9	Sequence 9, App1
41	99.5	3.8	600	1	US-08-374-155A-10	Sequence 10, App1
42	99.5	3.8	600	2	US-08-785-396-10	Sequence 10, App1
43	99	3.8	629	1	US-08-374-155A-4	Sequence 4, App1
44	99	3.8	629	2	US-08-785-396-4	Sequence 4, App1
45	99	3.8	974	3	US-08-938-291A-4	Sequence 4, App1

## ALIGNMENTS

RESULT 1  
US-09-041-718-5  
; Sequence 5, Application US/09041718A  
; Patent No. 6225075  
; GENERAL INFORMATION:  
; APPLICANT: Bard, Martin  
; TITLE OF INVENTION: DNA encoding sterol methyltransferase  
; FILE REFERENCE: 740.003051  
; CURRENT APPLICATION NUMBER: US/09/041,718A  
; CURRENT FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-041-718-5

Query Match 8.0%; Score 207; DB 4; Length 363;  
Best Local Similarity 28.6%; Pred. No. 3,9e-12;  
Matches 68; Conservative 38; Mismatches 88; Indels 44; Gaps 8;

OY 280 KLDLPGKVLVDGCGIGGGDFVYAEVYDVEYVIGDLSIMWIS--FALERSIGLKAVEF 337  
DB 114 QLELPGKVLVDGCGIGGDFVYAEVYDVEYVIGDLSIMWIS--FALERSIGLKAVEF 173  
OY 338 EVADCTKPDYNSFDVYASDITLHIDRPALEFRSHKMLPGKVLISDYC-KSAGTP 396  
DB 174 VKADFMKMPSDNFDVAVYAEVYDVEYVIGDLSIMWIS--FALERSIGLKAVEF 233  
OY 397 SAEPAVIR---ORGYLDHVKAYGKML---KDAEVEVYI-----AENR----- 434  
DB 234 NNATHKRIKDEIEIGNGLPIRSTRQCLQAVKDAGF-EVIMDKDLADSDPLPWLPLDPS 292  
OY 435 -----TDQFLOYQKELDLAEQKDFIDFSPEDVNDIVDGKAKTL 476  
DB 293 RFSLSRFLTTGVRITIRNMVKVL--EYVGADPGRSQRVSSFLKAAEGALVEGKKET 348

RESULT 2  
US-09-036-987A-13  
; Sequence 13, Application US/09036987A  
; Patent No. 6143526  
; GENERAL INFORMATION:  
; APPLICANT: Baltz, Richard H.  
; APPLICANT: Broughton, Mary C.  
; APPLICANT: Crawford, Kathryn P.  
; APPLICANT: Madhuri, Krishnamurthy  
; APPLICANT: Metito, Donald J.

```

; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes for Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-036-987A-13

Query Match
Best Local Similarity 25.7%; Pred. No.2e-11;
Matches 53; Conservative 44; Mismatches 97; Indels 12; Gaps 5;

QY 272 ETTKEFYSKLDLPGOKVLDVGGGIGGDFYMAENDVEVVGIDLSINMISFA--LERSI 329
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 38 QLDLFDLKAALRRGALFLDLCGCGNGOPVYRACASGVRTGTVNNOHIAATRLANET 97
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 330 GLCAVEFEVADCTKKDYSPNSFDVYSRTIILHIDOKPALFRSFHKWLKPGKVLISDY 389
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 98 GLAGSLEFDLVDAQOLPYRPGFOAMAMOSVVOIVDOAAIREVHRILEPGRFVLGDI 157
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 390 CKSAGTPSAEFAAYIRRGYDLHVKAYGKMLDAGF----VEVIAENR--TDQFIQVL 442
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 158 ITFVRLPE--EYAAV--WTGTAHTLNSFTALVSEAGEILELVLDLNTQTCMWSVYDEL 214
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 443 QKELDALDEKDFIDDESEEDYNDI 468
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 215 LRLKIDELAGVEPAVGYQORVLGDI 240
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 3
US-09-029-603-2
; Sequence 2, Application US/09029603
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bietenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Pospiech, Andreas
; TITLE OF INVENTION: staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-20555/A/PCT
; CURRENT APPLICATION NUMBER: US/09/029,603
; CURRENT FILING DATE: 1998-03-20

```

```

; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Streptomyces longisporoflavus
; FEATURE:
; OTHER INFORMATION: methyl transferase-like protein
; US-09-029-603-2

Query Match
Best Local Similarity 24.3%; Score 179; DB 4; Length 280;
Matches 49; Conservative 37; Mismatches 72; Indels 44; Gaps 6;

QY 257 YERN-----FGP---GYVSTGGLFTTKE-----FVSKIDLPKGGKVLDDVC 294
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 18 YDRLTISAMNDGSFNPVNHIGYWDTPGSEATIEBAMDRLTDVFERLNAYATSHVLDLC 77
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 295 GIGGDFYMAENDVEVVGIDLSINMISFA--LERSIGLCAVEFEVADCTKKDYSPNSF 352
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 78 GVGGPGLRVARTGARVTGISISEQIRTNRLAAEAGVADRAVFGHGDAMKLPFADASF 137
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 353 DVIYSRDTILHIDOKPALFRSFHKWLKPGKVLISD-----YCSAGT 395
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 138 DAYMALESICHPMDROOVTEFVGVRLRPGGRIVLTDLFFENHPKAVRHRIIDFCDDLMS 197
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 396 PSA--EFAAYIRRGYDLHDV 414
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 198 TTADIDYVALLHRSGLRLREI 219
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 4
US-08-457-245-9
; Sequence 9, Application US/08457245
; Patent No. 5573915
; GENERAL INFORMATION:
; APPLICANT: BARRY III, Clifton E.
; APPLICANT: YUAN, Ying
; TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA INVOLVED
; TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC ACIDS IN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,245
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chambers, Guy W.
; REGISTRATION NUMBER: 30,617
; REFERENCE/DOCKET NUMBER: 15280-216000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant

```

## RESULT 6

GENERAL INFORMATION:

```

: APPLICANT: Bard, Martin
: TITLE OF INVENTION: DNA encoding sterol methyltransferase
: FILE REFERENCE: 740.003US1
: CURRENT APPLICATION NUMBER: US/09/041,718A
: CURRENT FILING DATE: 1998-03-13
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 361
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: US-09-041-718-4

```

Query Match	6.28;	Score 160;	DB 4;	length 361;
Best Local Similarity	23.68;	Pred. No. 1.6e-07;		
Matches 59;	Conservative 41;	Mismatches 110;	Indels 40;	Gaps 9;

```

Qy 213 GAYVKSKKNNQISLW-----QVDSDDKGFQRFLLSSQYFNSILERYV 260
      | : : : | : : : | : : :
Db 39 GGSLSAEKVQDNKYQWSFFRRPKELTAEKVPDVFDTFNLVTDIYEWGQSFH---- 94

```

```

Qy      261  FGPPIVSTGLETTK---EYVSKLDLPGQKVLVDGCGIGGPFYMAENVDEYVGCIDL 316
      | : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      95  FSPSPGKSHKDATRLHEMAVDLIQVPGQKILVDGCGVGGPMRAIASHRANVGCITI 154

```

```

Oy 317 SINNISFAL--ERSIGLKCAVEFEVADCTKKDPENSEFOVYISRTILHLI0DKPALFRSE 374
      : :: | : | | : : : : | | | : :
Db 155 NEYOVRARLHNKKAGLDALCEVVCNFTQMPFEDNSFGGAYSTATCHAPKLEEVYAEI 214

```

```

Oy 375 HKWLKPGKVLISDYCKSAGCIPPSAEFA---YIR-----QRGVDLHDVKA---GKMLK 4222
    :: ||| | | : : | : || | | : |
Db 215 YRVLKPG-----SMYSYEWVYITEKFAEDDEHAEVIOGIERGDALPGLRAIVYDIAETAK 269

```

Qy	423	DAGFEVIAE	432
		::	
Db	270	KVGF-EIVKE	278

RESULT 8  
US-09-041-718-2  
; Sequence 2, Application US/09041718A

; APPLICANT: Bard, Martin  
 ; TITLE OF INVENTION: DNA encoding sterol methyltransferase  
 ; FILE REFERENCE: 740.003051

```

; NUMBER OF SEQ ID NOS: 9
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 2

```

```

: TYPE: PRT
:
: ORGANISM: Candida albicans
US-09-041-718-2

```

Query Match	5.7%	Score 148.5;	DB 4;	Length 376;
Best Local Similarity	24.0%;	Pred. No. 2.3e-06;		
Matches 63; Conservative	42;	Mismatches 102;	Indels 55;	Gaps

Db 53 WDGISKDEE-KRLNDYSQTLHHYNLVTDPEYEGWSSFHFSRY-KGEAFQATARR 110

[illegible]

```

225  TGNCAVLELVAVLCCINNVLEINSEFVLIAGVLIINLTVCFALDLSGFHNVNLFQGVNVLDS  267
      | : : : : : | | : : : : : | : : : : : |
Db  167 KYHNDKLSIVKGDQMDFESESFAVVALENTVHAPVLEGVSEIYKVLKPGG-----  221

```

```

OY 388 DYCSAGRPSMFPAY--IRQRGD---LHDYKAYCKMLKDAQFVIVLENRFDQIOY 441
Db 222 -----IFGVLEVWMTDKDIDENNEHRKTAIGLEWOD-GIPKMYSRKVAEQALKN 269
OY 442 LOKELDALEOEKDDFIDDFSEE 463
Db 270 VGFE--IEYQKD--LADVDE 286

```

RESULT 9  
US-08-935-263-16  
; Sequence 16, Application US/08935263A

```
; APPLICANT: Furuichi, Yasuhiro
; APPLICANT: Hoshino, Tatsuo
; APPLICANT: Kimura, Hiroshi
; APPLICANT: Kimura, Hiroshi
```

```

; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
; FILE REFERENCE: Biotin Genes
; CURRENT APPLICATION NUMBER: US/08/935,263A

```

```

; EARLIER FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1

```

```

; TYPE: PRT
;
; ORGANISM: Kurthia sp.
US-08-935-263-16

```

Query Match	4.9%;	Score 128.5;	DB 3;	Length 276;
Best Local Similarity	22.9%;	Pred. No. 0.00013;		
Matches 63: Conservative	45:	Mismatches 97:	Indels 77:	Gaps

```

QY 11 EKEFEFKYIWEHSVDIVTEAMMLDSQSDLDKKEVEYSMLPEYEG---SVLEBAGD 6/
    ::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 DKQLSKRFESEH-----AKTYDAYA-NVQKNMAQLVDLLPQKNSQRINILEIGCGT 54

```

```

      | | : : | : : : | | : : | :
Db  55 GYLTRLVNTFPNASITAVDIAFGMEVAKGTIME-DRVTFLLCADIEEMTLN---ENYDL 110
      | | : : | : : : | | : : | :
136 TEGNMTIHWI CDEENIEDI VEDMT YGT KDCGVTEDBESCEUOSCDUKRK SMDTTPVDEBDEY 185

```

Db 111 IISNATFQWLN--LPGTIEQLTRLPEGNLIF-----STFG 146

QY 186 TKIFECHMDDSGNSYELSLGCKCIGAYVSKKNQNSLWQKVDSEDDKGFQRELD 245

QY 246 SSQYKENSILRYERVPGYVSTGGLETTKEFYSKLDLKPQKV 289  
|| :| :||| :||:| |

RESULT 10  
US-08-457-245-8

: PATENT NO. 557512  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: BARRY III, Clifton E.  
 :  
 : APPLICANT: YUAN, Ying

TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS  
 TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC A  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourile and Crew



APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899  
FILING DATE: 04-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-52517  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
TELEFAX: (619)238-0062  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1084 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-455-543A-56

Query Match 4.6%; Score 120; DB 1; Length 1084;  
Best Local Similarity 19.6%; Pred. No. 0.0075;  
Matches 94; Conservative 76; Mismatches 158; Indels 152; Gaps 26;

77 KASOVIALDFIESYIK-----KNESINGHYKNVFKMC-----ADVTSPS 115  
533 KSGEPVLDLDALENDIVEIRKMKIDSGEGRITLVKSDERYIDKGNITYWTP 592  
116 LNISNSVDIIFSNMLWYSDEEVERLVE-RMLKMKP-----GGYIFP-RESC----- 163  
593 VNGIDVSLAVLPYISFYIKAKLEETITQARYSETLKPNEESGYTFIAPRDYCDLK 652  
164 -----FHOSGDKRKSNTPHYRPRFYTKIKKECHMODDSNSY---ELSLIGC 209  
653 ISDNTEFLNFEEDRKTPNPNNSCNAD--LNRVLLDAGFTNELVONTWSKOKNITGV 710  
210 KC-----IGAVYKSKKNOQISMLMK-VSEDDKGFQRFUDSSQYKFNSTLRERY 260  
711 KAFRVYTDGIDTRYKPKAEEN-----WQENPEYEDSFYKRSIDNNYFTA--PFFNK 763  
261 FPGQVYSTGL--ETTKFVSKLDKPGQVLDVCGIGGDEYMAENYVEVVDLSTI 318  
764 SGRGAYESGIMVSKAVEIYIGKLLKPA-----VYGIKIDV 799  
319 N-WISFALERSIGLKAVEFEVADCTKDYPENSDVITYSDTILHIOQKPALEFRSHKW 377  
800 NSWIENFTKTSINDPCA--GPVCDCKR-----NS-DVM--DCVI----- 833  
378 LKPGGKVLIS---DYCSAG-----TPS-----AEFAAYIRORGYDLHDYKAYKMLKD 423  
834 LDDGGFLMANHDDYTNQIRFGELDPSLMRHLVNISYVAFKNSYIOYSCGEGAPKQ 893  
424 -----AGFEVIAE-----NRTDOFIQVL--OKELDALBOEKDDFIDFSEE 463  
894 GAGHRSAYVSVAIDIIQIGMWATPAAMWSIIQOFLSLTFPRLLEAVEMEDDDFAASLKQ 953

RESULT 13

US-08-223-305C-56  
Sequence 56, Application US/08223305C  
Patent No. 5851824  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/223,305C  
FILING DATE: April 4, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/868,354  
FILING DATE: April 10, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
TELEFAX: (619)238-0062  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1084 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-223-305C-56

Query Match 4.6%; Score 120; DB 2; Length 1084;  
Best Local Similarity 19.6%; Pred. No. 0.0075;  
Matches 94; Conservative 76; Mismatches 158; Indels 152; Gaps 26;

77 KASOVIALDFIESYIK-----KNESINGHYKNVFKMC-----ADVTSPS 115



```

Db 533 KSQEPYIDFLDAELNDIKYKIRNKMTDGSGEKTFRTLVKSODERYIDKGNRTYTWP 592
QY 116 LNISPNSVDIIFSWMILMYLSDSEVERLVE-RMLKWLKP-----GGYIFF--RESC----- 163
Db 593 VNGIDYSLALVLPYSPFYIAKLEETITQARYSETLKPDPNFEESGYFFIAPRODCNDLK 652
QY 164 -----FHOSGDHKRSNPTHYREPPRYKIRFECHMODDSGNSY---ELSLIGC 209
Db 653 ISDNTEFLNFMNEFIDKTPNPNCSNAD--LINRVLLDAGFTNELVONWYSKOKNKG 710
QY 210 KC-----IGAVYKSKKNOISWLMOK--VDSDDKGFQFRLDSSQYKFNSTIRYERV 260
Db 711 KARFVYDGGITRYYPKAGEN-----WQENPEYEDSFYKRSLDNDNYVFTA--PYFNK 763
QY 261 FGPQVSTGCL--ETTKFVSKLDLKPQOKVLDVCGIGGDFYMAENVYDVEVVGIDLSI 318
Db 764 SGPGAYESGIMVSKAVEIYIOGKLKPA-----VVGIRKIDV 799
QY 319 N-MISFALERSIGLKCAVEFEVADCTKKDYPENSEFDVYISRDITLIHIDOKALFRSPFKW 377
Db 800 NSWIENFTKTSIRBPCL--GPVCDCKR-----NS-DVM--DCVI----- 833
QY 378 LKPGKVLIS---DYCKSAG-----TPS-----AEFAAYIRORGYDLHDYKAYGKMLKD 423
Db 834 LDDGFFILMANHDDYTNQIGRFEGEIDPSLMRHLVNISYVAFNKSYSYQSCPEGAPKQ 893
QY 424 -----AGFVEVIAE-----NRTDOFIQVL--QKELDALQEKDQDFIDDFSEE 463
Db 894 GAGHRSAVPSVADILQIGMWATAAAMSILQOFLSLTFPRLLEAVEMEDDDFTASLSKO 953

```

```

RESULT 14
US-08-455-543A-53
: Sequence 53, Application US/08455543A
: Patent No. 5792846
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: METHOD OF INVENTION: METHODS
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-2926
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: Fastseq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,543A
: FILING DATE: May 31, 1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/223,305
: FILING DATE: April 4, 1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/868,354
: FILING DATE: April 10, 1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/620,250
: FILING DATE: 30-NOV-1990
: PRIOR APPLICATION DATA:

```

```

: APPLICATION NUMBER: US 07/482,384
: FILING DATE: 20-FEB-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/603,751
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US89/01408
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/176,899
: FILING DATE: 04-APR-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-52517
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619)238-0999
: TELEFAX: (619)238-0062
: INFORMATION FOR SEQ ID NO: 53:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1103 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: internal
US-08-455-543A-53

```

```

Query Match 4.6%; Score 120; DB 1; Length 1103;
Best Local Similarity 19.6%; Pred. No. 0.0077;
Matches 94; Conservative 76; Mismatches 158; Indels 152; Gaps 26;

```

```

QY 77 KASQVIALDFIESYK-----KNESINGHYKNKKEMC-----ADVSPS 115
Db 552 KSQEPYIDFLDAELNDIKYKIRNKMTDGSGEKTFRTLVKSODERYIDKGNRTYTWP 611
QY 116 LNISPNSVDIIFSWMILMYLSDSEVERLVE-RMLKWLKP-----GGYIFF--RESC----- 163
Db 612 VNGIDYSLALVLPYSPFYIAKLEETITQARYSETLKPDPNFEESGYFFIAPRODCNDLK 671
QY 164 -----FHOSGDHKRSNPTHYREPPRYKIRFECHMODDSGNSY---ELSLIGC 209
Db 672 ISDNTEFLNFMNEFIDKTPNPNCSNAD--LINRVLLDAGFTNELVONWYSKOKNKG 729
QY 210 KC-----IGAVYKSKKNOISWLMOK--VDSDDKGFQFRLDSSQYKFNSTIRYERV 260
Db 730 KARFVYDGGITRYYPKAGEN-----WQENPEYEDSFYKRSLDNDNYVFTA--PYFNK 782
QY 261 FGPQVSTGCL--ETTKFVSKLDLKPQOKVLDVCGIGGDFYMAENVYDVEVVGIDLSI 318
Db 783 SGPGAYESGIMVSKAVEIYIOGKLKPA-----VVGIRKIDV 818
QY 319 N-MISFALERSIGLKCAVEFEVADCTKKDYPENSEFDVYISRDITLIHIDOKALFRSPFKW 377
Db 819 NSWIENFTKTSIRBPCL--GPVCDCKR-----NS-DVM--DCVI----- 852
QY 378 LKPGKVLIS---DYCKSAG-----TPS-----AEFAAYIRORGYDLHDYKAYGKMLKD 423
Db 853 LDDGFFILMANHDDYTNQIGRFEGEIDPSLMRHLVNISYVAFNKSYSYQSCPEGAPKQ 912
QY 424 -----AGFVEVIAE-----NRTDOFIQVL--QKELDALQEKDQDFIDDFSEE 463
Db 913 GAGHRSAVPSVADILQIGMWATAAAMSILQOFLSLTFPRLLEAVEMEDDDFTASLSKO 972

```

```

RESULT 15
US-08-223-305C-53
: Sequence 53, Application US/08223305C
: Patent No. 5851824
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven

```

```

APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: Mcclain, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/866,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1103 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-223-305C-53

```

Query Match 4.6%; Score 120; DB 2; Length 1103;  
 Best Local Similarity 19.6%; Pred. No. 0.0077;  
 Matches 94; Conservative 76; Mismatches 158; Indels 152; Gaps 26;

```

QY 77 KASQVIALDFESYIK-----KNESINGHYKKNVKEMC-----ADVTSPS 115
DB 552 KSQPVLDLFDLADLEIKVEIKNMKIDSGEKTEFTLVKSODERYIDKGNFTYTWP 611
QY 116 LNIIPNSVVDIFFSNWLMYLSDEEVERLVE-RMLKWLKP-----GGYIFF--RESC----- 163
DB 612 VNGIDYSLAVLVPYTSYYIKAKLEBITTQARYSETLKPDPNFESGYTFTIAPRDYCDL 671

```

```

QY 164 -----FHSGDHRKRSNPTHYREPRFYKIFRECHMODDSGNSY---ELSLIGC 209
DB 672 ISDNNTTEFLNPNFEIDRKTTPNPSQND--LINRVLLDAGFTNELVQNYMSKQKNIKGV 729
QY 210 KC-----IGATYKSKKNOQISWLMQK-VSEDDKGFORFLDSSQYKNSILRTERV 260
DB 730 KAREVVDGGITRYVPKKEAGEN-----WOENPETYEDSFYRSISLDNDYVETA--PYFNK 782
QY 261 FGPQVSTGGL--ETTKFVSKLDLKPQGVLDVCGGIGGSDFYMAENYDVEVYIDLSI 318
DB 783 SGPGAYESGIMYSKAVEIYIGKILKPA-----VGIKIDV 818
QY 319 N-MISFALERSIGKCAVEFEVADCTKDPENSFDVLYSDTLIHLQDKPALERSFHKW 377
DB 819 NSWIEINFKTSTRDPCA--GPVCDCKR-----NS-DVM--DCVI----- 852
QY 378 LKPGKVLIS--DYCKSAG-----TPS-----AEFAAYIRQGYDLHDVKAYGKMLKD 423
DB 853 LDDGGFLMANHDDYTNQIGRFGSEIDPSLMRHLVNIISVAFNKSVDYQSVCEPGAAPKO 912
QY 424 -----AGFEVETIAE-----NRTDQFIQVL--QKELDALQEKDDFTIDFSEE 463
DB 913 GAGHRSAVPSVADILQIGMWATAAAMSILQOFLSLFPRLLAEVEMEDDFTASLSKQ 972

```

Search completed: September 20, 2001, 16:19:43  
 Job time: 181 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:17:22 ; Search time 17.12 Seconds  
(without alignments)  
2198.027 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598  
Sequence: 1 MAASAMGYLOEREVEFKKYWI.....KLVRTGEGHOGLFIARKM 494.

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2086	80.3	555	2 H96762	hypothetical protein
2	1287.5	49.6	374	2 F96525	protein T1N15.23 l
3	498.5	19.2	437	2 T29330	hypothetical prote
4	449.5	17.3	495	2 T27936	hypothetical prote
5	207	8.0	353	2 T06795	probable 24-sterol
6	204.5	7.9	317	2 S76618	hypothetical prote
7	191	7.4	367	2 T06780	probable 24-sterol
8	189.5	7.3	387	2 A81324	cyclopropane-fatty
9	187.5	7.2	227	2 C71246	hypothetical prote
10	185	7.1	418	2 S76226	hypothetical prote
11	183.5	7.0	344	2 B75306	probable cycloprop
12	182	6.8	344	2 T01572	24-sterol C-methyl
13	177	6.8	344	2 T04138	gamma-tocopherol m
14	176.5	6.8	306	2 S18533	eryg protein - Sac
15	176	6.8	382	2 A44292	cyclopropane-fatty
16	175.5	6.8	382	2 F85773	cyclopropane-fatty
17	175	6.8	283	2 T44579	steroid biosynthesi
18	174.5	6.7	227	2 C75210	24-sterol C-methyl
19	173	6.6	346	2 T10173	cyclopropane-fatty
20	171	6.6	389	2 H64571	24-sterol C-methyl
21	170.5	6.6	283	2 J05531	avermectin B 5-O-m
22	170	6.5	383	1 S42003	24-sterol C-methyl
23	166	6.4	389	2 B71865	cyclopropane-fatty
24	165.5	6.4	366	2 T33885	hypothetical prote
25	163.5	6.3	254	2 B72316	conserved hypotet
26	162.5	6.3	378	2 F84657	hypothetical prote
27	161.5	6.2	378	2 T39579	sterol methyltrans
28	160	6.2	361	2 S63686	24-sterol C-methyl

30	159.5	6.1	286	2 B70614	probable mmaA1 pro
31	159.5	6.1	391	2 T08338	conserved hypotet
32	157.5	6.1	269	2 C75569	probable methyltra
33	156.5	6.0	253	2 F83511	hypothetical prote
34	156.5	6.0	324	2 T42375	probable 24-sterol
35	156.5	6.0	394	2 C82951	conserved hypotet
36	156	6.0	379	2 T50969	probable DELTA(24)
37	154.5	5.9	278	2 T38024	hypothetical prote
38	151	5.8	273	2 T34740	hypothetical prote
39	151	5.8	352	2 T03845	probable 24-sterol
40	150.5	5.8	218	2 H84364	membrane protein l
41	150.5	5.8	262	2 S62195	hypothetical prote
42	150	5.8	776	2 B70797	probable transiera
43	149.5	5.8	256	2 F85505	unknown (imported)
44	149	5.7	245	2 B86738	hypothetical prote
45	149	5.7	251	2 A86071	hypothetical prote

## ALIGNMENTS

## RESULT 1

H96762  
hypothetical protein F6D5.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: H96762

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: H96762

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-555 <STO>

A:Cross-references: GB:AE005173; NID:q10092368; PIDN:AA012776.1; GSPDB:GN00141

A:Gene: F6D5.1

A:Map position: 1

Query Match 80.3%; Score 2086; DB 2; Length 555;

Best local similarity 77.7%; Pred. No. 7.1e-139; Mismatches 42; Indels 0; Gaps 0;

Matches 376; Conservative 66; Mismatches 42; Indels 0; Gaps 0;

QY	10	QERVPFKWYEHSDVLTVEAMMLDSQASDLDKREPEVLSMLPPYGGKSVLEAGIGR	69
DB	71	BERIOLKMKHSHVGVLSVEAMMLDSKASDLDKREPEVLSMLPPYGGKSVLEAGIGR	130
QY	70	FTGLAERASQVIALDFIESYIKKNESINGHYKKNVEMKCAVTSPTSINISNVDIIFSN	129
DB	131	FTTELAKAGQVIAVDFIESYIKKNENINGHYKKNVEMKCAVTSPTSINISNVDIIFSN	190
QY	130	WLWYLSDEVEYERVMKMLKPGGYIFFRSCFPHSGDHKKRSNPHYREPFYKIF	189
DB	191	WLWYLSDEVEYERVMKMLKPGGYIFFRSCFPHSGDHKKRSNPHYREPFYKIF	250
QY	190	KECHMODSGNSYELSLIGCKICGAYVKSKNQNIEMLMQKVSSEDDKGFORLDSOY	249
DB	251	KECHMODSGNSYELSLIGCKICGAYVKSKNQNIEMLMQKVSSEDDKGFORLDSOY	310
QY	250	KFNLSILYERVFPGYVSTGGLETTKEFVSRLDKPGQKVLVDCGIGGDFYMAENYDV	309
DB	311	KSSSILYERVFPGYVSTGGLETTKEFVSRLDKPGQKVLVDCGIGGDFYMAENYDV	370
QY	310	EVVIGIDLSINMKIPALERSIGLKCAVEEVAADCTKKDY PENSFVITYSRDTILHIQDKPA	369











Db	155	VKADPMKPPEDFDNFDAVYATEACHPADDPVGCCKEYRLVKPQCGAVVEWMTCTDHYDP	214
QY	397	SAEFAATYR---QNGYDLDHYKATGAKML--KDAGEYEV-----IAENR-----	434
Db	215	NNATHRKIKDEIELGNGLPDIRSTROCIQAKDQAGFEVMDKDLDAEDSPLPWYLPDPSR	274
QY	435	-----TDQFIQVLCKELDALEQEDDFELDFSEEDYNDIYDGGKAKL	476
Db	275	FSLSFRLTISGVRMITRPMVAL--EYVGLAPQSSSEVSNFLKAAAGCLVEGGCKEI	329

**RESULT 14**

C96673  
gamma-tocopherol methyltransferase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96673  
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chlu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzalli,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shlita, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141, MUID:21016719  
A:Accession: C96673  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-348 <STO>  
A:Cross-references: GB:AB05173; NID:98099780; PIDD:AAD38271.2; GSPDB:GM00141  
C:Genetics:  
A:Gene: F13011.27  
A:Map position: 1

Query Match	6.88	Score 177	DB 2	Length 348
Best Local Similarity	27.68	Pred. No. 4.9e-05		
Matches 54	Conservative 32	Mismatches 62	Indels 48	Gaps 8

```

QY      264 GYVSTGLETTFKFNYSKLDLRFQKVLVDYCGGIGGDFPMAENYEVWGIDLSIMNISF 323
          | : | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      119 -----IDEEEEKI-----KKVVDVCGGIGGSSRLASKFCAECIGTILSPYQAKR 164

```

```

QY      324 A--LERSIGLKAVEFEVADCSKKDDYBNSFDVIYSRDITLHIDQKRALFRSFHKWYLRG 381
      | | : | | | : : : | | : : : | |
Db      165 ANDLAAQSLAHKASQVADALDQRFEEGKFDLWMSMEGEMHPDKAKFYKELVRYAARG 224

```

```
QY      382 GKVLISDYCK---SAG 394
        |:::| :| |||
Db      225 GRITVTWCHRNLSAG 240
```

RESULT 15

eryG protein. *Saccharopolyspora erythraea*  
C:Species: *Saccharopolyspora erythraea*  
C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 22-Oct-1999  
C:Accession: SI8533; SI6747  
R:Haydock, S.F.; Dowson, J.A.; Dhillon, N.; Roberts, G.A.; Cortes, J.; Leadlay, P.F.  
MOL. Gen. Genet. 230, 120-128, 1991  
A:Title: Cloning and sequence analysis of genes involved in erythromycin biosynthesis in  
methylintrasterases.  
A:Reference number: SI8530; MUID:92079886  
A:Accession: SI8533  
A:Molecule type: DNA

A:Residues:1-306 <HA>  
A:Cross-references: EMBL:X60379; NID:g48941; PIDN:CAA42929.1; PID:g581650  
C:Genetics:  
A:Gene: cryg  
A:Start codon: CNG  
C:Superfamily: bioc homology  
C:Superfamily: bioc homology <BIOC>  
E:82-18/Domain: bioc homology

Query Match	6.8%	Score 176.5	DB 2	Length 306
Best Local Similarity	27.8%	Pred. No. 4.5e-05		
Matches 60	Conservative 36	Mismatches 81	Indels 39	Gaps 9

Qy 316 LSIINMISFALERS--IGLKCAVEFEVADCTKKDPENSEFDVYISKDTILHIQDKPALERS 373  
| : : : | | : : : | : | : : | : : | :  
Db 113 LTPSHVRIASERAERENVQDRQFKESATDLPGAEFTDRTVSPESALHYEPRTDFEKG 172

QY 374 FHKMLKPGGKVLISDYCK-----SAGTPSAEFAAYINRGYDLH-----DYKAY 417  
: |||| : | | | | : |  
Db 173 AFEEVLKPGGVLAIGDIIPLDIREPGSDGPPK----LAPQRSGLSGGIPVENWVPRETY 227

```
QY 418 GKMLKDAFVEV-----IAENRTDQFIQVLQKE 445
      |:::|::|:|      :|:::|::|
Db 228 AKQLRAGGFVDVEVKSVRDNNVMEPWLDYWLRLKIQDE 263
```

Search completed: September 20, 2001, 16:20:09  
Job time: 167 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 20, 2001, 16:19:23 ; Search time 14.71 Seconds  
(without alignments)  
1150.388 Million cell updates/sec

Title: US-09-525-885-2  
Perfect score: 2598  
Sequence: 1 MAASAMGYLOEREFKKYWI.....KLVRTGEEOGWLFIAKKM 494

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues  
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	176	6.8	381	1	CFA_ECOLI
2	170	6.5	382	1	ERGB_YEAST
3	161.5	6.2	378	1	ERGB_SCHPO
4	148.5	5.7	376	1	ERGB_CANAL
5	144	5.5	244	1	YXBB_BACSU
6	144	5.5	251	1	UBIE_ECOLI
7	135.5	5.2	345	1	YTJ7_STRFR
8	135	5.2	224	1	MEMH_BACST
9	134.5	5.2	232	1	GRGB_LACLA
10	128	4.9	240	1	UBIG_ECOLI
11	128	4.9	417	1	YAT1_SYNP6
12	127.5	4.9	209	1	UBIG_SALTY
13	126	4.8	212	1	Y829_SYNY3
14	123.5	4.8	197	1	TEHB_ECOLI
15	123	4.7	251	1	Y095_HAEIN
16	120	4.6	1021	1	Y912_HAEIN
17	119	4.6	1021	1	DPOK_NEUCR
18	118.5	4.6	247	1	Y0EM_BACSU
19	117.5	4.5	220	1	PIMT_PYRHO
20	117.5	4.5	394	1	YLR3_PSEBU
21	117	4.5	255	1	BIOC_SERMA
22	116.5	4.5	286	1	COO3_RAT
23	115	4.4	287	1	CEFA_MYCTU
24	115	4.4	302	1	CEFA_MYCTU
25	114.5	4.4	1091	1	CIC2_HUMAN
26	113.5	4.4	1106	1	CIC2_RABIT
27	112	4.3	250	1	Y028_RICPR
28	111	4.3	3068	1	POLG_PEMVC
29	110.5	4.3	1435	1	DP43_MYCPU
30	110	4.2	243	1	YX42_MYCTU
31	110	4.2	1091	1	CIC2_RAT
32	109.5	4.2	626	1	HTB6_BACSU
33	109.5	4.2	699	1	HS81_ARATH

34	106.5	4.1	251	1	BIOC_ERWHE	006898 erwinia her
35	106.5	4.1	1675	1	POL_RTBVP	P27502 rice tungro
36	106	4.1	261	1	TAM_MYCTU	053698 mycobacteri
37	106	4.1	621	1	HTPG_HELPY	P56116 helicobacte
38	105.5	4.1	621	1	HTPG_HELPY	092mm2 helicobacte
39	105	4.0	715	1	ZP2_CANFA	P47983 canis fami1
40	105	4.0	1071	1	VATA_YEAST	P17255 saccharomyc
41	104.5	4.0	203	1	PMTA_RHOSH	005197 rhodobacter
42	104.5	4.0	366	1	Y906_METJA	058316 methanococc
43	104.5	4.0	1294	1	RRPO_MCMVM	P09498 white clove
44	103.5	4.0	1700	1	HS81_ARATH	P27323 arabidopsis
45	103.5	4.0	1376	1	RPOB_RICTY	P77941 ticketelia

## ALIGNMENTS

RESULT 1  
CFA\_ECOLI STANDARD; PRT; 381 AA.  
ID P30010;  
AC 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE (EC 2.1.1.79)  
GN CPA OR CDPA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 1-8, AND CHARACTERIZATION.  
RC SRRATN-K12;  
RX MEDLINE=93075691; PubMed=1445840;  
RA Wang A.-Y., Grogan D.W., Cronan J.E. Jr.;  
RT "Cyclopropane fatty acid synthase of Escherichia coli: deduced amino acid sequence, purification, and studies of the enzyme active site.";  
RL Biochemistry 31:11020-11028(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SRRATN-K12 / RR28;  
RA Eberhardt S.M.R., Richter G., Gimbel W., Werner T., Bacher A.;  
RN Submitted (NOV-1992) to the EMBL/Genbank/DBJ databases.  
RX [3]  
RP SEQUENCE FROM N.A.  
RC SRRATN-K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SRRATN-K12;  
RX MEDLINE=97251357; PubMed=9097039;  
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
RA Itoh T., Kasai K., Kashimoto K., Kimura S., Kitahara M.,  
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,  
RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,  
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubraman S.,  
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,  
RA Yamamoto Y., Horiuchi T.;  
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map.";  
RL DNA Res. 3:363-377(1996).  
RN [5]  
RP FUNCTION: TRANSFERS A METHYLENE GROUP FROM S-ADENOSYL-L-METHIONINE TO THE CIS DOUBLE BOND OF AN UNSATURATED FATTY ACID CHAIN  
CC RESULTING IN THE REPLACEMENT OF THE DOUBLE BOND WITH A METHYLENE BRIDGE.  
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + PHOSPHOLIPID

```

CC CCLEFINIC FATTY ACID = S-ADENOSYL-L-HOMOCYSTEINE + PHOSPHOLIPID
CC CYCLOPROpane FATTY ACID.
CC
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M98330; AAA23562.1; -.
CC DR EMBL; X69193; -. NOT ANNOTATED_CDS.
CC DR EMBL; AE000261; AAC74733.1; -.
CC DR EMBL; D90809; BAA15428.1; -.
CC DR EMBL; D90810; BAA15437.1; -.
CC DR PIR; A44292; A44292.
CC DR ECGene; EG11531; cfa.
CC Pfam; PF00972; Flav_N55; 1.
CC KW Transferase; Methyltransferase; Lipid synthesis.
CC FT INT_METH 0
CC FT CONFLICT 1 S -> R (IN REF. 2).
CC FT CONFLICT 7 E -> G (IN REF. 2).
CC FT CONFLICT 24 S -> N (IN REF. 2).
CC FT CONFLICT 38 I -> T (IN REF. 2).
CC FT SEQUENCE 381 AA; 43777 MW; 1R07B20C7E08ADF CRC64;
CC
CC Query Match 6.8%; Score 176; DB 1; Length 381;
CC Best Local Similarity 29.3%; Pred. No. 1,2e-05;
CC Matches 65; Conservative 27; Mismatches 80; Indels 50; Gaps 12;
OY 184 FYTKIFK-----ECHNODDSGNSYEISLIGCKIGAYVSKKNQNOISMLMQRVSED 236
Db 74 FFSKVLRRGLENOQLPHNRKD-----TLRIGARLRFN--LQSKRR-----AWYVKR--EHY 119
OY 237 DKG---FORFLDSSOYKFNLSIRRYERVPGYVSTGLET*-----KEPVSKLDIKPGQK 288
Db 120 DLGNDLFFSRMLD-----PMPQYSCAY---WKDADNLESAQQAKLKICEKTLQIKPMR 169
OY 289 VLDVCGSGISGGDEYVAENYDEYVGVISIDISIMNISFALERSIGLKCAVEFAVDCKRKDYP 348
Db 170 VLDIGCGGGLAHYVANSYVDYSVGVTTISADQQAQAEKRCESLDVTTILQ-----DYR 222
OY 349 E--NSFDVIYSRDTLLIQDK--PALFRSFFKWKLPKGGKVL 386
Db 223 DLNDQFDRIYVSGMEHVGPKNYDYPFAVVDNRNLKPEGIFLL 264
RESULT 2
ERG6_YEAST STANDARD; PRT; 382 AA.
AC P25087;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DEFIL(24)-STEROL C-METHYLTRANSFERASE (EC 2.1.1.41).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94262330; PubMed=8203167;
RA Hardwick K.G., Pelham H.R.B.;
RT "SBD6 is identical to ERG6, and encodes a putative methyltransferase
RT required for ergosterol synthesis."
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=94312403; PubMed=8038180;
RX

```

RA Melthinda A.A., Beavis A.D., Trumbley R.J.;  
RT "mutations in LIST (ERG6) gene confer increased sodium and lithium  
RN uptake in *Saccharomyces cerevisiae*";  
RL Biochim. Biophys. Acta 1193:107-117(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA Gentles S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 1-257 FROM N.A.  
RX MEDLINE-91285426: PubMed-2060792;  
RA Hussain M., Lenard J.;  
RT "Characterization of PDR4, a *Saccharomyces cerevisiae* gene that  
RT confers pleiotropic drug resistance in high-copy number: identity  
RT with YAP1, encoding a transcriptional activator.";  
RN Gene 101:149-152(1991).  
RN [5]  
RP ACETYLATION.  
RA Garets J.I., Fitcher B., Kobayashi R., Latter G.I., Schwender B.,  
RL Volpe T., Warner J.R., McLaughlin C.S.;  
RA Submitted (SEP-1994) to the SWISS-PROT data bank.  
CC -1- FUNCTIONAL ACTIVITY: METHYLTRANSFERASE REQUIRED FOR ENOGENOL SYNTHESIS.  
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 5-ALPHA-CHOLESTA-  
CC 8,24-DIEN-3-BETA-OL -> S-ADENOSYL-L-HOMOCTEINE + 24-METHYLENE-  
CC 5-ALPHA-CHOLEST-8-EN-3-BETA-OL.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X74249: CAA52308.1; -;  
DR EMBL: S72460: AAB31378.1; -;  
DR EMBL: Z49810: CAA89944.1; -;  
DR EMBL: X53850: CAA37826.1; -;  
DR PIR: PEO402: PEO402.  
DR PIR: S17001: S17001.  
DR PIR: S35982: S35982.  
DR PIR: SA2003: SA2003.  
DR SWISS-2DPAGE: P25087; YEAST.  
DR YEPD: 6410; -;  
DR SGD: S0004467: ERG6.  
KW Sterol biosynthesis; Transferase; Methyltransferase; Acetylation.  
FT INTR MEM 0  
FT MOD RES 1  
FT CONFLICT 379 379 ACETYLATION  
SQ SEQUENCE 382 AA: 43299 MW; D/6BA9E3DD9CD71B CRC64;

```

Query Match          6.5%: Score 170; DB 1; Length 382;
Best Local Similarity 23.6%: Pred. No. 3,2e-05;
Matches      65; Conservative    44; Mismatches 127; Indels   40; Gaps
QY      190 KECMHODSGNSYEISLIGCCICAGYVASKNNO---ISMUKVSSEDDKG-----239
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      14 RELH-GDIDIGKGTGJAL-----MSKNNSAKOEAVQKYLLNMWGRTDKDAEERLE 63
QY      240 -----FORELDISOYKENSLILRYERVEPGYSTGTGLETTKEFVS-KLDLPQGV 289
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      64 DYENATHSYWVVDFEFYEGMGSSGFHFSR-PYKKEESFAASIAHEHLAKRAGIORDDL 122
QY      290 LDVGCGIGGGDEPYMAENVADVWVGIDLSINNISR*-LERISIGLKAAVEEVADCTKKDY 347
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      123 LDVGCYGVGSPAREIRLFQGCVAITGLNNNDIYLAAKAYTAKKLYNLSDMDVVKGDGFKMDF 182
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      348 PENSFDVIYSRDTLIHIODKPALFRSFHKWLKPGCAVLISDYCKS---AGTPSAEEPAAY 403
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      183 EENTEDKYYAIEATCCHAPKLEGVYSEIKTYLKPGRGTAYAEVWMTDXYDENNPBHRKIAY 242

```



Db 270 VGFE---LEYOKD--LADVDDE 286

RESULT 5

ID YXBB\_BACSU STANDARD; PRT; 244 AA.

AC P46326;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE HYPOTHEICAL 28.2 KDA PROTEIN IN ALDX-ASNH INTERGENIC REGION.

GN YXBB OR VETCR.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / BGSC1A1;

RX MEDLINE=96093926; PubMed=7584049;

RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;

RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis genome between the gnt and lol operons.";

RL DNA Res. 2:61-69(1993).

CC -1- SIMILARITY: TO E.COLI YAFE.

CC -1- SIMILARITY: TO METHYLTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AB005554; BAA21596.1; -

DR EMBL: Z99124; CAB16025.1; -

DR Sublist; BG11352; yxsb.

DR Interpro: IPR000339; -

DR Pfam: PF01209; Uble.methyltran; 1.

KW Hypothetical protein.

SO SEQUENCE 244 AA; 28222 MW; 5B008371AFD3D02B CRC64;

Query Match 5.5%; Score 144; DB 1; Length 244;

Best Local Similarity 20.4%; Pred. No. 0.0013;

Matches 55; Conservative 49; Mismatches 118; Indels 48; Gaps 8;

OY 232 VSEDDKGFQFPLDSSQYKFNSTL--RYERVFQGVYSTGLETTFKFSKLDKPGQKV 289

DB 2 ITAEECTYKEFVLYOSYLYPLGARLSRMYSF-----KMKVI 40

OY 290 LDVGGGIGGDDFYAENVYDVEVIGDLSINMISFALF--RSIGLKAVEFEVADCTKKDY 347

DB 41 IDMGTEGGYISTOLAKTNAHVAVDINPMHETAEAKKSGVSLISFDLEDVHHLST 100

OY 348 PENSFDVYSRDTLTIQDPALEFRSFHMKLPGKVLISDYCSACT-----PSA 398

DB 101 ADQYADRTVYSCLHMHEDVVKGLKECYRVLAQGLIVLDTFPMQSHLEIMKQIKEP 160

OY 399 EFAYAIR--QRYDLHDVAYGKMLKDAQFEVYIEN--RTDQFLQVLOKELDALEQ 452

DB 161 EYRFVAREAEESYSEFDIQF--VODAGIPNYSLETFFHFLPEDFTESDELDADPLWE 217

OY 453 KDDFIDFSEEDYNDIVDGKAKLVRTTEG 482

DB 218 QNDOSTDHEITESVT-----MLTTEKEKG 242

RESULT 6

ID UBIE\_ECOLI STANDARD; PRT; 251 AA.

AC P27851;

DT 01-AUG-1992 (Rel. 23, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE UBIOQUINONE/MENAOQUINONE BIOSYNTHESIS METHYLTRANSFERASE UBIE

DE (EC 2.1.1.-).

GN UBIE.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=92358234; PubMed=1379743;

RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;

RT "Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 to 86.5 minutes.";

RL Science 257:771-778(1992).

RN [2]

RP REVISION TO 101.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP CHARACTERIZATION, AND VARIANT UBIE401.

RX MEDLINE=97197541; PubMed=9045837;

RA Lee P.T., Hsu A.Y., Ha H.T., Clarke C.F.;

RT "A C-methyltransferase involved in both ubiquinone and menaquinone biosynthesis: isolation and identification of the Escherichia coli ubiE gene.";

RL J. Bacteriol. 179:1748-1754(1997).

CC -1- FUNCTION: CONVERTS DMOH2 INTO DMOH2 AND DMKH2 INTO MKH2.

CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 2-POLYPRENYL-6-METHOXY-1,4-BENZOQUINOL = S-ADENOSYL-L-HOMOCYSTEINE + 2- POLYPRENYL-3-METHYL-6-METHOXY-1,4-BENZOQUINOL.

CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DDMETHYLMENAOQUINOL = S-ADENOSYL-L-HOMOCYSTEINE + MENAOQUINOL.

CC -1- PATHWAY: INVOLVED IN BOTH UBIOQUINONE AND MENAOQUINONE BIOSYNTHESIS.

CC -1- SIMILARITY: BELONGS TO THE UBIE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M87049; AAA67628.1; -

DR EMBL: AB000459; AAC76636.1; -

DR PIR: S30722; S30722.

DR EcoGene: EG11473; ubiE.

DR Interpro: IPR000339; -

DR Pfam: PF01209; Uble.methyltran; 1.

DR PROSITE: PS01183; UBIE\_1; 1.

DR PROSITE: PS01184; UBIE\_2; 1.

KW Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase; Methyltransferase.

FT VARIANT 142 142 G -> D (IN UBIE401 ALLELE, DEFECTIVE).

FT CONFLICT 101 101 P -> L (IN REF. 1).

SO SEQUENCE 251 AA; 28057 MW; 646F3E09A0075A9E CRC64;

Query Match 5.5%; Score 144; DB 1; Length 251;

Best Local Similarity 26.8%; Pred. No. 0.0014;

Matches 52; Conservative 36; Mismatches 88; Indels 18; Gaps 7;

OY 283 LKPGQKVLVGGIGG--GDFYMAENVYDVEVIGDLSINMISFALF--RSIGLKAVEFE 338



```

Db      61 VRRCQTVLADLAGTGLTAAFSRLVGETGKRVYLADINESMPEKMGREKLNRNIGVANEYV 120
      339 VADCTKRDYFENSFDVYISRDITLHIDOKPALFSPFKMLKPGKVLISYCSAGTSSA 398
      121 QANNEALPFEDNTFDCTTISFGLRNWTDKOKALRSKRVLKPGGRLLVLEFSKPIEPLS 180
      399 EFAYAIRQGYDLHDVAKAYKML-KDAGFVEVIAENRTDQFIQVLOKELDALEQEKDFTI 457
      181 K-AV---DAVSFHLPRISLVANADSYRYLAESIR-----MHPDDQTLKAMQD-- 227
      458 DDFSEEDYNDIVDG 471
      228 AGESVDYVNLTAG 241

RESULT 7
YT37_STRFR STANDARD; PRT; 345 AA.
ID YT37_STRFR
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE HYPOHETICAL 37.1 KDA PROTEIN IN TRANSPOSON TN4556.
OS Streptomyces fradiae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1906;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-TN4556;
RX MEDLINE-90185236; PubMed-2155856;
RA Siemieniak D.R., Slightom J.L., Chung S.T.;
RT "Nucleotide sequence of Streptomyces fradiae transposable element
   TN4556: a class-II transposon related to Tn3."
RL Gene 86:1-9(1990).
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M29297; AAA8564.1; -.
CC PIR: J00429; J00429.
CC InterPro: IPR000339; -.
CC Pfam: PF01209; Uble_methyltran; 1.
CC Hypothetical protein; Transferase; Methyltransferase;
CC Transposable element.
CC KW
CC SQUENCE 345 AA; 37113 MW; 027CDA5B3F8AA7CC CRC64;

Query Match 5.2%; Score 135.5; DB 1; Length 345;
Best Local Similarity 26.9%; Pred. No. 0.009;
Matches 45; Conservative 22; Mismatches 67; Indels 33; Gaps 6;

Db      281 LDLPGRKVLDDVGGGIGG--GDFYMAENVDEYVGGIDLSIMISFALERSGLCAVEFE 338
      122 LDAPRGESALDGGCGPGTDGLTAKAVSPGSGVIGIDSSQEMVQARRRENLP-AVEVE 180
      339 VADCTKRDYFENSFDVYISRDITLHIDOKPALFSPFKMLKPGKVLISYCSAGTSSA 398
      181 LQDHTHTPLEGSGIDCATRTDYLQHVADPQAQALAEARRVLRPGSRVLVWGEPDMDSLTIDY 240
      390 CKSAGTSPSAEFAAYIRQGYDLHDV-----KAYGKMLKDAGF 426
      241 -----PDLEYS-----RAYTRHTVDKIVRNGVIGRQALRLADAGF 276

RESULT 8

```

```

MENH_BACST STANDARD; PRT; 234 AA.
ID MENH_BACST
AC O86169;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2-HEPTAPRENYL-1,4-NAPHTHOQUINONE METHYLTRANSFERASE (EC 2.1.1.-)
DE (MENAQUINONE BIOSYNTHESIS METHYLTRANSFERASE).
GN MENH OR MENG.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN-ATCC 10149;
RX MEDLINE-97284706; PubMed-9139683;
RA Koike-Takeshita A., Koyama T., Ogura K.;
RT "Identification of a novel gene cluster participating in menaquinone
   (vitamin K2) biosynthesis. Cloning and sequence determination of the
   RT 2-heptaprenyl-1,4-naphthoquinone methyltransferase gene of Bacillus
   stearothermophilus."
RL J. Biol. Chem. 272:12380-12383(1997).
CC -1- FUNCTION: CONVERTS DMK42 INTO MK42.
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DEMETHYLMENAQUINOL =
CC S-ADENOSYL-L-HOMOCYSTEINE + MENAQUINOL.
CC -1- PATHWAY: LAST STEP IN MENAQUINONE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE UBIE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D87054; BAA32500.1; -.
CC InterPro: IPR000339; -.
CC Pfam: PF01209; Uble_methyltran; 1.
CC PROSITE: PS01184; UBIE_2; FALSE_NEG.
CC Menaquinone biosynthesis; Transferase; Methyltransferase.
CC KW
CC SQUENCE 234 AA; 27128 MW; 8C1CDA06A525F953 CRC64;

Query Match 5.2%; Score 135; DB 1; Length 234;
Best Local Similarity 21.9%; Pred. No. 0.0057;
Matches 47; Conservative 42; Mismatches 80; Indels 46; Gaps 7;

Db      246 SSOY-KFNSILRYRVRGPGYVSTGLETTEKFEYSKDLKPGQKVLDDVGGGIGGDFYMA 304
      17 SAHYDRMNSVISFR-----HLKMRKDVRRNRNVQKKGKALDVCGGTADWTITALA 66
      305 ENTDFE--VVGIDLSIMISFALE--RSIGLKAVEFEVADCTKRDYFENSFDVYISRD 360
      67 EAVGPECKYVGLPSEMMKLKGVQKVARGLH-NVKLIHGNAMOLPPDNSFDVYITGFG 125
      361 ILLHIDOKPALFSPFKMLKPGKVLISDYCKSAGT-
      126 LRNPVDYMTVLKEMHAKRTKPGITV-----CLETSQPLFLFGRQRLYYFRFIMPLFKLL 181
      397 -SAEFAAYIRQGYDLHDVAKAYKMLKDAGFVEV 429
      182 AKSYEEYSWLGESARERPPGRDELAEMFRAGFVDV 216

RESULT 9
GRCB_LACIA STANDARD; PRT; 252 AA.
ID GRCB_LACIA
AC P49016;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

```

DT 30-MAY-2000 (Rel. 39, last annotation update)  
DE PROBABLE MENAQUINONE BIOSYNTHESIS METHYLTRANSFERASE (EC 2.1.1.-)  
OS (GERC2 PROTEIN HOMOLOG)  
OC Lactococcus lactis (subsp. lactis) (Streptococcus lactis)  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC2;  
RX MEDLINE=93374846; PubMed=8366036;  
RA Geller B.L., Ivey R.G., Trempe J.E., Hettlinger-Smith B.;  
RT "Cloning of a chromosomal gene required for phage infection of  
RT Lactococcus lactis subsp. lactis C2.";  
RT J. Bacteriol. 175:5510-5519(1993)  
CC  
CC -1- FUNCTION: CONVERTS DMK2 INTO MKH2.  
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DEMETHYLMENAQUINOL =  
CC S-ADENOSYL-L-HOMOCYSTEINE + MENAQUINOL.  
CC -1- PATHWAY: LAST STEP IN MENAQUINONE BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE UBIE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: L14679; AAA03166.1; -  
DR InterPro: IPR000339; -  
DR Pfam: PF01209; Ubie\_methyltran; 1.  
DR PROSITE: PS01183; UBIE.1; 1.  
DR PROSITE: PS01184; UBIE.2; 1.  
KM Menaquinone biosynthesis; Transferase; Methyltransferase.  
SQ SEQUENCE 252 AA; 28449 MW; 7D971E3A4BD804ED CRC64;

Query Match 5.2%; Score 134.5; DB 1; Length 252;  
Best Local Similarity 21.5%; Pred. No. 0.0069;  
Matches 52; Conservative 44; Mismatches 97; Indels 49; Gaps 7;

QY 215 YVMS--KKNOQISWLMQKXDESDKGFQRFLLSDSSYKFNISLRYEFGPGYSTGGL 272  
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
DB 3 YNMSMTVNERVEQELFNSISDSD-----KMAAISFKO-----HDLW 41

QY 273 TTKEFVSKLKGQKVLVDGCGGSGGFYMAENY--DVEVGDLSINMISFLER-SI 329  
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
DB 42 RAKTKRMGDL-TGLSLDLCGGTGTDFLSDSVSGKVIQGLDFSENMELAKAKLKE 100

QY 330 GLKCAVEFEVADCTKKDYPENSPDYISRDITLHIDKPALFRSFHMLKPGKVL---- 385  
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
DB 101 EAKKNIEFLGNMALPFEKSGFYVITGYGLRNTPLYLVKIEFVLKPGGVVCIET 160

QY 386 -----ISDYCKSAGTSPSAEPFAYIRQGYDLHDVYKAYGKMLKDAGEV 427  
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
				: : : : :	: : : : :	: : : : :	: : : : :
DB 161 SHPTLPYKQAFELYKRNVMPLFGKVFPAKSLKEYQWLQKSAEDPDKATLEELFRKAGFV 220

QY 428 EV 429  
DB 221 AV 222

RESULT 10  
UBIG\_ECOLI STANDARD; PRT; 240 AA.  
AC P17993; P76924;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, last sequence update)  
DT 01-OCT-2000 (Rel. 40, last annotation update)  
DE 3-DEMETHYLUBIQUINONE-9-3-METHYLTRANSFERASE (EC 2.1.1.64)  
DE (3,4-DIHYDROXY-5-HEXAPRENYLBENZOATE METHYLTRANSFERASE) (DHMB  
DE METHYLTRANSFERASE).

GN UBIG OR PUPF.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=93123968; PubMed=1479344;  
RA Wu G., Williams H.D., Zamanian M., Gibson F., Poole R.K.;  
RT "Isolation and characterization of Escherichia coli mutants affected  
RT in aerobic respiration: the cloning and nucleotide sequence of ubig.  
RT Identification of an S-adenosylmethionine-binding motif in protein,  
RT RNA, and small-molecule methyltransferases.";  
RT J. Gen. Microbiol. 138:2101-2112(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OV6;  
RX MEDLINE=88201664; PubMed=2834621;  
RA Hussain K., Elliott E.J., Salmood G.P.C.;  
RT "The parD-mutant of Escherichia coli also carries a gyrAa mutation.  
RT Mol. Microbiol. 1:259-273(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RT Science 277:1453-1474(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RA Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiuchi T.,  
RA Ikemoto K., Inada T., Isono K., Itoh T., Kanai K., Kasai H.,  
RA Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,  
RA Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,  
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y.,  
RA Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE +  
CC 3-DEMETHYLUBIQUINONE-9 - S-ADENOSYL-L-HOMOCYSTEINE +  
CC UBIQUINONE-9.  
CC  
CC -1- PATHWAY: UBIQUINONE BIOSYNTHESIS.  
CC -1- SUBUNIT: HOMODIMER (PROBABLE).  
CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: Y00544; CA468610.1; -  
DR EMBL: M87509; AAA24714.1; -  
DR EMBL: AE000313; AAC75292.1; -  
DR EMBL: D90854; BAA16049.1; -  
DR EMBL: D90855; BAA16051.1; -  
DR PIR: S03757; S03757.  
DR PIR: A47682; A47682.  
DR Ecogene: EG1143; ubig.  
DR InterPro: IPR000339; -  
DR Pfam: PF01209; Ubie\_methyltran; 1.  
KM Ubiquinone biosynthesis; Transferase; Methyltransferase.  
SQ SEQUENCE 240 AA; 26555 MW; D4EB4707A06F5613 CRC64;

Query Match 4.9%; Score 128; DB 1; Length 240;

Best Local Similarity 24.8%; Pred. No. 0.019;  
Matches 59; Conservative 38; Mismatches 75; Indels 66; Gaps 13;

```

OY 217 KSKKNONISLMQKVDSEDKGFQ---RFLDSSQYKFNILYERFEGP-----263
DB 5 KSPVNHN-----VHEETAKFEAVASRWMD-----LEGEFKLHINPLRL 45
OY 264 GYVS--TGLETTKEFYSKLDLKPQKVLVYGCIGGDFYMAENVY---VVGIDLSI 318
DB 46 GYIERRAGLE-----GKRVLDVGC---GGG--ILASSMAREGATVGLDMGF 88
OY 319 NMISFALERSIGLKCAVEFEVADCTKKDYP---ENSEVDIYSRDTILHIQDPALFRSFH 375
DB 89 EPILOVA--KLHALSFGIOVDVVOETVEEHAHAKHAGQYDVVTCMELHEVDPQSVVACA 146
OY 376 KWLPGGKVLISDYCKSA---GTPSAEFAAYIRORGYDLDHVKAYGKMLKDAGFEV 428
DB 147 QLVKPGGDFVFTLNRRGKSWLMAVGAELYILRMVPG--THDVKKFKIPALLGWD 202

```

```

RESULT 11
YATL_SYNP6
ID YATL_SYNP6 STANDARD; PRT; 417 AA.
AC P08442;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ATP SYNTHASE SUBUNITS REGION ORF 1.
OS Synecococcus sp. (strain PCC 6301). (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1139;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87311713; PubMed=3041005;
RA Cozens A.L.; Walker J.E.;
RT "The organization and sequence of the genes for ATP synthase subunits
in the cyanobacterium Synecococcus 6301. Support for an
endosymbiotic origin of Chloroplasts."
RL J. Mol. Biol. 194:359-383(1987).
CC -1- SIMILARITY: TO M.TUBERCULOSIS RV2067C.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X05302; CAA28921.1; -
DR PIR; S07286; S07286.
KM Hypothetical protein.
SQ SEQUENCE 417 AA; 47531 MW; 4F96D4F5E7C0F99B CRC64;

```

```

Query Match 4.9%; Score 128; DB 1; Length 417;
Best Local Similarity 25.3%; Pred. No. 0.041;
Matches 55; Conservative 29; Mismatches 87; Indels 46; Gaps 8;
OY 288 KVLVDVGGIGGDFYMAE--NVDVEVVGIDLSINMISFALER--SIGLKCAVEFEVADCTK 344
DB 74 RILDAGGCTGVSTDYLAHLNPSAETIADISAGTLAAVQEQCRSSGVAARLHFQOLSLYD 133
OY 345 KDYPENSFDVIYSRDTILHIQDPALFRSFHKLKPGKVLISDYCKSAGTPSAEFAAYI 404
DB 134 VAQIPGDFDQNCVGLHLEDPDRGLAALASKLAPGILHIFYAELG-----RAEI 186
OY 405 RQ-----RGYDLHDVKAYGKMLKDAGFEVIAENFTDOFIQVLRQKDALDQEK 453
DB 187 ROMQEAIALLOGERRGRDVGVAIGREI---FSQLPANNR---LRRREERWALEENOR 238
OY 454 DDFIDDF-----SEEDYNDIYDGMKAKILVRTTEGQOW 486

```

DB 239 DECFADMYVHPQELDYN-----TETLRW 262

```

RESULT 12
UBIG_SALTY
ID UBIG_SALTY STANDARD; PRT; 209 AA.
AC P37431;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 3-DEMETHYLBIQUINONE-9-3-METHYLTRANSFERASE (EC 2.1.1.64)
DE (3,4-DIHYDROXY-5-HEXAPENYLENZOATE METHYLTRANSFERASE) (DHMB
DE METHYLTRANSFERASE) (FRAGMENT).
GN UBIG.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=94253018; PubMed=8195103;
RA Jordan A., Gilbert I., Barde J.;
RT "Cloning and sequencing of the genes from Salmonella typhimurium
RT encoding a new bacterial ribonucleotide reductase."
RL J. Bacteriol. 176:3420-3427(1994).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE +
CC 3-DEMETHYLBIQUINONE-9 -> S-ADENOSYL-L-HOMOCYSTEINE +
CC BIQUINONE-9.
CC -1- PATHWAY: BIQUINONE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE UBIG/COO3 FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X72948; CAA51451.1; -
DR PIR; S32628; S32628.
DR SycGene; SG10405; ubiG.
KM Ubiquinone biosynthesis; Transferase; Methyltransferase.
FT NON TER
SQ SEQUENCE 209 AA; 23061 MW; 3D07EB2B31D0D5E7 CRC64;

```

```

Query Match 4.9%; Score 127.5; DB 1; Length 209;
Best Local Similarity 26.4%; Pred. No. 0.017;
Matches 48; Conservative 34; Mismatches 63; Indels 37; Gaps 10;
OY 264 GYVS--TGLETTKEFYSKLDLKPQKVLVYGCIGGDFYMAENVY---VVGIDLSI 318
DB 13 GYIERRAGLE-----GKRVLDVGC---GGG--ILASSMAREGATVGLDMGF 55
OY 319 NMISFALERSIGLKCAVEFEVADCTKKDYP---ENSEVDIYSRDTILHIQDPALFRSFH 375
DB 56 EPILOVA--KLHALSFGIEVEYVOETVEEHAHAKHAGQYDVVTCMELHEVDPQSVVACA 113
OY 376 KWLPGGKVLISDYCKSA---GTPSAEFAAYIRORGYDLDHVKAYGKMLKDAGFEV 430
DB 114 QLVKPGGDFVFTLNRRGKSWLMAVGAELYILRMVPG--THDVK---KFKIPALLSWV 168
OY 431 AE 432
DB 169 DE 170

```

```

RESULT 13
Y829_SYNP3
ID Y829_SYNP3 STANDARD; PRT; 212 AA.

```

```

AC 055423;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE PUTATIVE METHYLTRANSFERASE SLI0829 (EC 2.1.1.-).
GN SLI0829.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
ON NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 648 to 928 of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D64003; BAI0522.1; -.
DR InterPro: IPR000339; -.
DR Pfam: PF01209; Ublc_methyltran; 1.
KM Hypothetical protein; Transferase; Methyltransferase.
SQ SEQUENCE 212 AA; 23071 MW; BAD77E2E605A7001 CRC64;

Query Match 4.88; Score 126; DB 1; Length 212;
Best Local Similarity 27.7%; Pred. No. 0.022;
Matches 54; Conservative 32; Mismatches 77; Indels 32; Gaps 12;

QY 266 VSTGGLFTTKEF-VSKIDLRGKQVLDVCGIGGDFPMANVDEVVGDLSINMISFA 324
DB 25 LNVGGEERFQDLPLENLAIISPGQVLDCCGGQATVYLADS-GATVGLDAS----- 76
QY 325 LENSIGLKCAVEFEVADCT-----KDDYP--ENSFDVYSHDTLIHQDKPALFRSF--- 374
DB 77 -PKALG---RAKINVPQATVYQGLAEDLPREGSEFDLVHT-SVALH-EMTPAQQLSIISG 130
QY 375 -HKWLKFGKVLISDYKSAG---TPSAEF-AAVIRQGTDLHDVKA YGKMLKDGAFVE 428
DB 131 VHRVLKPGGIFALVDLRHPSNMLFMPPLAIFMGLFETETAMOLINTD-LGSLIDQAGFTV 189
QY 429 VIAENRTDFIOVLQ 443
DB 190 VRKHLVAGGSLOVLIQ 204

RESULT 14
TEHB_ECOLI
ID TEHB_ECOLI STANDARD; PRT; 197 AA.
AC P25357; P76866;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE TELLURITE RESISTANCE PROTEIN TEHB.
GN TEHB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=91285417; PubMed=2060788;

```

```

RA Walter E.G., Weiner J.H., Taylor D.E.;
RT "Nucleotide sequence and overexpression of the tellurite-resistance
RT determinant from the IncHI1 plasmid pH1508a.";
RL Gene 101:1-7(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.O.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshtima T., Saito N., Sampei G., Seki Y., Sivasubaram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28 0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [4]
RP GENE MAPPING.
RC STRAIN=K12;
RX MEDLINE=94222856; PubMed=8169225;
RA Taylor D.E., Hou Y., Turner R.J., Weiner J.H.;
RT "Location of a potassium tellurite resistance operon (tehA tehB)
RT within the terminus of Escherichia coli K-12.";
RL J. Bacteriol. 176:2740-2742(1994).
CC -1- FUNCTION: RESPONSIBLE FOR POTASSIUM TELLURITE RESISTANCE WHEN
CC PRESENT IN HIGH COPY NUMBER, PROBABLY BY INCREASING THE REDUCTION
CC RATE OF TELLURITE TO METALLIC TELLURIUM WITHIN THE BACTERIUM.
CC OTHERWISE, PHENOTYPICALLY SILENT.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE PLASMID ENCODED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M74072; AAA19564.1; -.
DR EMBL: AE000240; AAC74512.1; -.
DR EMBL: D90782; BAA15059.1; -.
DR EMBL: D90783; BAA15064.1; -.
DR PIR: JQ1018; JQ1018.
DR Ecogene: EG11884; tehB.
KM Antibioc resistance: Tellurium resistance.
SQ SEQUENCE 197 AA; 22531 MW; BF0D2D64F999C2E8 CRC64;

Query Match 4.88; Score 123.5; DB 1; Length 197;
Best Local Similarity 28.6%; Pred. No. 0.031;
Matches 59; Conservative 29; Mismatches 83; Indels 35; Gaps 13;

QY 254 ILRYRYFGGVSTGGLETTKEFVSKLD-LKPGQAVLDVCGIGGDFYMAEN-YDVEY 311
DB 2 IIRDENYFTDKYEITL--RTSHSEVLKAVKVKPG-KTLDLGCCGNGRNSIYLAANGYDVA 57
QY 312 VGIDLSINMISFA-LER--SIGLKCAVEFEVADCTKKDYENSFDVYSHDTLIHQDK- 367
DB 58 -----WDKNAMSIANVERIKSIENLDLHTRVVDLNLNLT-F-DROYDFLSTVLMFLEAKT 112

```

QY	DB	QY	DB
368	-PAFRRSHKMLKGGKGLI-----SDYCSACTPSAEFAAYTRQ--RGV-----	409	
113	IPGLIANNQRCRTKGGVNLIIYAAMDPAADYPCVTGFPFAFKGELRRYYEGMERKYNNEDV	173	
410	-DLHDVKA YGKMLKRDAGFVEVIAENR	434	
173	GELHRTDANGNRIR-LRFATYLAIRKK	197	



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:18:13 ; Search time 19.82 seconds  
(without alignments)  
3297.613 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598  
Sequence: 1 MAASAMGVLOREVEFKKYYI.....KIVRTGEQOQGLFIKKM 494

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL\_16:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_protent:\*

12: sp\_unclassified:\*

13: sp\_vertebrate:\*

14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2598	100.0	494	10 Q9M571	Q9M571 spinacia ol
2	2105.5	81.0	491	10 Q9ER44	Q9ER44 arabisidopsis
3	2093	80.6	498	10 Q9LVH3	Q9LVH3 arabisidopsis
4	1287.5	49.6	374	10 Q9LP63	Q9LP63 arabisidopsis
5	767.5	29.5	180	10 Q9LP64	Q9LP64 arabisidopsis
6	498.5	19.2	437	5 Q22993	Q22993 caenorhabd1
7	449.5	17.3	495	5 Q23552	Q23552 caenorhabd1
8	223	8.6	565	2 Q9KJ20	Q9KJ20 actinopolys
9	207	8.0	363	10 Q41586	Q41586 tritlicum ae
10	207	8.0	363	10 Q41587	Q41587 tritlicum ae
11	204.5	7.9	317	2 Q55809	Q55809 streptomyc
12	199	7.7	278	2 Q9EY12	Q9EY12 streptomyc
13	196	7.5	346	10 Q82720	Q82720 nicotiana t
14	193.5	7.4	279	2 Q9KJ21	Q9KJ21 ectochloro
15	191	7.4	367	10 Q43445	Q43445 glycine max
16	189.5	7.3	275	2 Q9X508	Q9X508 streptomyc
17	189.5	7.3	283	2 Q9X509	Q9X509 streptomyc
18	189.5	7.3	387	2 Q9PMB2	Q9PMB2 campylobact
19	189	7.3	272	2 Q52570	Q52570 amycolatops

20	187.5	7.2	227	1	Q57965	Q57965 pyrococcus
21	187	7.2	336	10	Q9LM02	Q9LM02 arabisidopsis
22	185	7.1	318	2	P74388	P74388 synechocyst
23	184	7.1	349	10	Q82434	Q82434 nicotiana t
24	184	7.1	438	2	Q9KZ58	Q9KZ58 streptomyc
25	183.5	7.1	462	2	Q9RSD7	Q9RSD7 delinococcus
26	182	7.0	344	10	Q49215	Q49215 zea mays (m
27	182	7.0	344	10	P93852	P93852 zea mays (m
28	179	6.9	278	2	Q9RNB8	Q9RNB8 microcystis
29	178	6.9	348	10	Q9ZSK1	Q9ZSK1 arabisidopsis
30	177	6.8	348	10	Q9XIP9	Q9XIP9 arabisidopsis
31	176.5	6.8	306	2	Q54095	Q54095 saccharopol
32	175.5	6.8	283	2	Q9S0N6	Q9S0N6 streptomyc
33	175.5	6.8	349	10	Q82426	Q82426 oryza sativ
34	174.5	6.7	227	1	Q9V268	Q9V268 pyrococcus
35	173	6.7	346	10	Q24328	Q24328 rickia com
36	171	6.6	389	2	Q25171	Q25171 helicobacte
37	169.5	6.5	244	2	Q9EZC1	Q9EZC1 staphylococ
38	166	6.4	389	2	Q9ZKG8	Q9ZKG8 helicobacte
39	165.5	6.4	366	5	Q9YXP1	Q9YXP1 caenorhabd1
40	163.5	6.3	254	2	Q9X036	Q9X036 thermotoga
41	162.5	6.3	581	10	Q64850	Q64850 arabisidopsis
42	160	6.2	361	10	Q39227	Q39227 arabisidopsis
43	160	6.2	361	10	Q9LN25	Q9LN25 arabisidopsis
44	159.5	6.1	286	2	P94922	P94922 mycobacteri
45	159.5	6.1	391	1	Q52025	Q52025 halobacteri

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	494 AA.
ID	Q9M571			
AC	Q9M571			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DE	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	PHOSPHOETHANOLAMINE N-METHYLTRANSFERASE (EC 2.1.1.103).			
GN	PEAMT.			
OS	Spinacia oleracea (Spinach).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;			
OC	Caryophyllales; Chenopodiaceae; Spinacia.			
OX	NCBI_TaxID=3562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=SPINACH LEAF;			
RX	MEDLINE=20261526; PubMed=10799484;			
RA	Nuccio M.L., Ziemak M.J., Henry S.A., Weretillyn E.A., Hanson A.D.;			
RT	"Phosphoethanolamine N-methyltransferase from spinach: cDNA cloning by			
RT	complementation in Schizosaccharomyces pombe and characterization of			
RT	the recombinant enzyme."			
RL	J. Biol. Chem. 275:14095-14101(2000).			
DR	EMBL: AF237633; AAF61950.1; -			
DR	InterPro: IPR000051; -			
DR	InterPro: IPR000780; -			
KW	InterPro: IPR001601; -			
KW	Transferase; Methyltransferase.			
SQ	SEQUENCE 494 AA; 56361 MW; 7F2537C8EAB8413B CRC64;			
Query Match	100.0%; Score 2598; DB 10; Length 494;			
Best Local Similarity	100.0%; Pred. No. 1.4e-168;			
Matches 494;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MAASAMGVLOREVEFKKYYIHSVDLTFEAMMLDSQASDLDKVERPEVLSMPPECKSV 60			
DB	1 MAASAMGVLOREVEFKKYYIHSVDLTFEAMMLDSQASDLDKVERPEVLSMPPECKSV 60			
QY	61 LEAGAGIGRTGTELAERASQVIALDFIESYIKKNSINGKXKVKFCAVYTSPLNISP 120			
DB	61 LEAGAGIGRTGTELAERASQVIALDFIESYIKKNSINGKXKVKFCAVYTSPLNISP 120			

QY	121	NSVDIIIFSWMILMYLSDEVEERLVERMLKMLKPGGIIFFRESCFHSGGHHKRSNPTHYR	180
QY	121	NSVDIIIFSWMILMYLSDEVEERLVERMLKMLKPGGIIFFRESCFHSGGHHKRSNPTHYR	180
QY	181	EPFRYTKIKPECHMODDSGNSYELSLGCKCGIYAAYKSKKNOISWLMQKVSDEDKCF	240
Db	181	EPFRYTKIKPECHMODDSGNSYELSLGCKCGIYAAYKSKKNOISWLMQKVSDEDKCF	240
QY	241	QRFLLDSOYKFNLSIRYERFPGPGYVSTGGLETTKEFVSKLDLPGQKVLVDVCGIGIGGD	300
Db	241	QRFLLDSOYKFNLSIRYERFPGPGYVSTGGLETTKEFVSKLDLPGQKVLVDVCGIGIGGD	300
QY	301	FYMAENYDEVYVIGIDLSTIMISFALERSIGLCAVFEVADCTKKDYPNSFDVYISRPT	360
Db	301	FYMAENYDEVYVIGIDLSTIMISFALERSIGLCAVFEVADCTKKDYPNSFDVYISRPT	360
QY	361	ILHIDDKPALFSPFHKWLKPGGKVLISDYCKSAGPSPSAFAAYITROGYDLHDVRYAGKM	420
Db	361	ILHIDDKPALFSPFHKWLKPGGKVLISDYCKSAGPSPSAFAAYITROGYDLHDVRYAGKM	420
QY	421	LKDAFVEVYIAENRDOEIOVLQKELDALDEQEKDFIDDSSEEDYNDIYDGKAKLVRTT	480
Db	421	LKDAFVEVYIAENRDOEIOVLQKELDALDEQEKDFIDDSSEEDYNDIYDGKAKLVRTT	480
QY	481	BGEQOQGLEFIARKM 494	
Db	481	BGEQOQGLEFIARKM 494	
RESULT 2			
Q9FR44	Q9FR44	PRELIMINARY:	PRT: 491 AA.
AC	Q9FR44:		
DT	01-MAR-2001 (TREMblrel. 16, Created)		
DT	01-MAR-2001 (TREMblrel. 16, Last sequence update)		
DT	01-MAR-2001 (TREMblrel. 16, Last annotation update)		
DE	SAM: PHOSPHO-ETHANOLAMINE N-METHYLTRANSFERASE.		
GN	NM1L		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;		
OC	Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_taxid=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. LANDSBERG ERECTA; TISSUE-STAGE TWO LEAVES;		
RX	MEDLINE=20567827; PubMed=11115895;		
RA	Biologisches C.P., McGraw P.;		
RT	"The Isolation and Characterization in Yeast of a Gene for Arabidopsis		
RL	S-adenosylmethionine:Phospho-Ethanolamine N-Methyltransferase.";		
RL	Plant Physiol. 124:1800-1813(2000).		
DR	EMBL; AF197940; AAC41121.1.		
DR	Transferase; Methyltransferase.		
Q0	SEQUENCE 491 AA: 56102 MW: DA4404EBBD3FA8D5 CRC64:		

	Query Match	81.0%	Score 2105.5;	DB 10;	Length 491;
	Best Local Similarity	78.9%;	Fred. No. 4e-135;		
	Matches 389;	Conservative 55;	Mismatches 46;	Indels 3;	Gaps 1.
Qy	1 MAASAMGVLOEVEEVKKRWIEHSDVLTAEAMMLDSQAOLDYEREPEVLSLPPYEKGSV	60			
	:     :     :     :     :     :     :     :     :     :				
Dd	1 MAAS---YEERDIDKNWIEHSADLYEAMMLDSRASLDKEERPEVLSLLPPIEGKSV	57			
Qy	61 LELGAGIRFTGELA EKASQVIALDFIESVIKKNESINGHYKNVKFMCAVDTSPLNTSP	120			
	:     :     :     :     :     :     :     :     :     :				
Dd	58 LELGAGIRFTGTGLAKRAGELITALDFIDIVIKKINSINGHYKNVKFMCADVTSPDLKITD	117			
Qy	121 NSVDIIFSNWLIMTYSDSEVERLVERMKLWKLPKGCIFFRESCFPQSGDHKKKSNPETHR	180			
	:  :				
Dd	118 GSLDIIFSNWLIMTYSDDEVELLRNMGWIGVGCIFFRESCFPQSGDKSKSNTETHR	177			

QY	181	BPREFTKJFKCHMOHDOSYEYELISJGCKIGAYVSKKNONISLMOKVSEDDKGF	24.0
QY	178	BPREFTKJFKCHMOHDOSYEYELISJGCKIGAYVSKKNONISLMOKVSEDDKGF	23.7
QY	241	QREFLDSSQYKFNISILRYERVFGPGYVSTGSGLETTKEFVSKLDPGOKVLDVGGGJGGD	30.0
Db	238	QREFLDSSQYKFNISILRYERVFGPGYVSTGSGLETTKEFVSKLDPGOKVLDVGGGJGGD	29.7
QY	301	FYMAENYVYEWVAVGJDLSTNMISFALERSIGLCAVEEVVADCTCKVDPENSPDYISRDT	36.0
Db	298	FYMAENYVYEWVAVGJDLSTNMISFALERSIGLCAVEEVVADCTCKVDPENSPDYISRDT	35.7
QY	361	ILIHIDKRALFRSEHKMLKPGGKYLISDYCKSAGTPSAEFAAYIRORGYLDHVKAYGKM	42.0
Db	358	ILIHIDKRALFRSEHKMLKPGGKYLISDYCKSAGTPSAEFAAYIRORGYLDHVKAYGKM	41.7
QY	421	LKDAGFVAVIENKRTDQFIYVLQKELDALDEKNDPDIIDSEEDYNDIYVGMKARVPTT	48.0
Db	418	LKDAGFVAVIENKRTDQFIYVLQKELDALDEKNDPDIIDSEEDYNDIYVGMKARVPTT	47.7
QY	481	EGEQOMGLFIANK 493	
Db	478	EGEQOMGLFIANK 490	

```

RESULT      3
09LVH3
ID 09LVH3          PRELIMINARY;          PRT;          498 AA.
AC 09LVH3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE METHYL TRANSFERASE-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tadata S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL: AB019230; BAR02720.1; -.
DR InterPro: IPR000051; -.
DR InterPro: IPR001601; -.
SQ SEQUENCE 498 AA; 56877 MW; E57693B915ECF0D8 CRC64;

```

Query Match	80.6%	Score 2093	DB 10	Length 498
Best Local Similarity	78.0%	Pred. No. 2.9e-134		
Matches 390	Conservative 54	Mismatches 46	Indels 10	Gaps 2
OY	1	MAASMGVLOEERFKKWIIEHSDVLVEAMMLDQASDLKVEREVLSTMLPRYEGKSV	60	
Db	1	MAAS---YEEERDQKNWIEHSADLTVEAMMLDSRADLDEERREVLSTLPRYEGKSV	57	
OY	61	LELGAGIGRFTGELAEKASQVIALDPFIESVKKNESINGHYNNVKFMCAADVTSPLSINISP	120	
Db	58	LELGAGIGRFTGELAQKGELIALDPIDIVAKKNSINGHYNNVKFMCAADVTSPLDKTTR	117	
OY	121	NSVDIIFFSNMLLMYLSDEEVERLVERMILKWLKPGGIIFRSCPFHOSGDHKRKSINPTHYR	180	



```

Db 118 GSDLIFFSNMLLMTLWSOXEVELLAEERNVGMKLVGGVIFIFRSCRHOGSDSKRSNPHYR 177
Oy 181 EPRFYTKI-----FKECHMODDSNGSYELSLIGCKICIGAYVSKKNQNOISMLQKVD 233
Db 178 EPRFYSKIDGETMKVFOECQTRDAAGNSFELSMIGCKICIGAYVKKKNQNOICIMQKVS 237
Oy 234 SEDDKGPORELDSSOYKKNSTLARKERPGPCVYSTGGLETTKEFVSKLDLKPQKVLVDG 293
Db 238 SENDGFGREFLDNVOYKSSGILTRERAVFGQGFYSTGGLETTKEFVSKMLKPQKVLVDG 297
Oy 294 CGIGGGDEYMAENADVEVWIGDLSINMSFPLERSIGLKCAFEFVADCTKKDPENSPFD 353
Db 298 CGIGGGDEYMAEKEDVHVYIGDLSVNMISFPLERAIGLSCVEFEVADCTTKHHPDMSFD 357
Oy 354 VIYSRDTLIHTIDOKPALERFSRHKMLKPGKVLISDYCKSACTPSAEAAVYIRORGYLHD 413
Db 358 VIYSRDTLIHTIDOKPALERFPFFMKLPGKVLISDYCKSRPPTPSAESEVYIKOGYLDHD 417
Oy 414 VKATGKMLKDGAFPEVVAENETDPIQVLYKELDALDEQKDDFIIDFSSDYDINDYGMK 473
Db 418 VQAGAGMLKDGAFEDVVAEDRTDQFMQVLRKELRVERKEKEFIIDFSKEDYDIDYGMK 477
Oy 474 AKLVRTTEGEQOOWGLFIATK 493
Db 478 SKLERCASDEQKGLFIATK 497

RESULT 4
O9LP63 PRELIMINARY; PRT; 374 AA.
AC O9LP63;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE TINI5.23.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxId=3702;
RN 11
RT 1
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC TINI5 from chromosome
1."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN 131
RT SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC020089; AAF79705.1; -.
DR InterPro: IPR0000051; -.
DR InterPro: IPR001601; -.

```

SEQ	SEQUENCE	374 AA;	42656 MW;	318D85E03110CC47	CNC64;
Query Match		49.6%;	Score 1287.5;	DB 10;	Length 374;
Best Local Similarity		66.4%;	Pred. No. 9.6e-80;		
Matches 243;		Conservative 36;	Mismatches 28;	Indels 59;	Gaps 2;
QY	187	KIFRCCHMODDSGNSYEISLICCKICIGAYVSKKNONQISWIMQKVDSEDDKGFQRFILDS	246		
DB	8	GVFQECQTRDASGNSFELSMVGCICIGAYVKNKKNONQICWIMQKVSVEENKDFQRFILDN	67		
QY	247	SOYKENSILIRYRVGPGVSGSL-----ETTRKE	276		
DB	68	VOYKSSGLIRYRVGEGVSTGFGNSILLSSYGHLYCLSVIMFLFSLTETTRKE	127		
QY	277	FVSKSLDKPGQVLDVVGCGIGGDPYMAENYDVEYVIGDLSINMISFALENSIGLKCAVE	336		
DB	128	FVAKKMDLKPQGVLDVVGCGIGGDPYMAENYVHVIGDLSINMISFALERIGLKCAVE	187		
QY	337	FEVADCTKDKDPENSFVDIYSHDTLLIHODKPALEFRSFHMKLPGKVLISDYCKSAGTP	396		
DB	188	FEVADCTKRTYTPDPSNFVDIYSHDTLLIHODKPALEFRFFKMLPGKVLITDYCRSAETP	247		
QY	397	SAEFAAYIRQRYDLDHDKAYGKMLKDGAFVETIAENRTDQFIOYLQKELDALEQKDF	456		
DB	248	SPEFAEYIKQRYDLDHDKAYGKMLKDGAFVETIAEDRTDQFIOYLQKELAEVKEKEEF	307		
QY	457	IDDFSE-----EDYNDIYDGKAKLVPTTEGEQMG	487		
DB	308	ISDFSEKTFELTTNSFEQNTCLCSHIDHKILLDQEDYNDIYDGKSAKLERTASGEQKMG	367		
QY	488	LEFIARK 493			
DB	368	LEFIADK 373			
RESULT	5				
Q9LP64	Q9LP64	PRELIMINARY;	PRT;	180 AA.	
AC	Q9LP64;				
DT	01-OCT-2000 (TREMBLrel. 15, Created)				
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)				
DT	TIN15.22.				
OC	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eurariopsis: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:				
OC	Magnoliophyta: eudicotyledons, core eudicots; Rosidae; eurosids II;				
OC	Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,				
RA	Shinn P., Altafi H., Bel Q., Chin C., Chio J., Choi E., Conn L.,				
RA	Conway A., Gonzales A., Hansen N., Hwang B., Koo T., Lam B., Lee J.,				
RA	Leuz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,				
RA	Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaverl A.,				
RA	Toriunli M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,				
RA	Ecker J.R.;				
RT	"Genomic sequence for Arabidopsis thaliana BAC TIN15 from chromosome				
RT	I.";				
RL	Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RA	Ecker J.R.;				
RL	Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.				
RL	[3]				
RP	SEQUENCE FROM N.A.				
RA	Ecker J.R.;				
RL	Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.				
RL	[4]				
RP	SEQUENCE FROM N.A.				
RA	Chen R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,				
RA	Khan S., Kim C., Altafi H., Bel B., Chin C., Chio J., Choi E.,				

RA Conn L., Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharly N.,  
 RA Nguyen M., Palm C., Plam P., Sakano H., Schwartz J., Southlack A.,  
 RA Thayer A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.,  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC020889; AAF79704.1; -  
 DR InterPro: IPR000051; -  
 DR InterPro: IPR000780; -  
 DR InterPro: IPR001601; -  
 SQ SEQUENCE 180 AA; 20472 MW; AAOFEF2DD27E556 CRC64;

Query Match 29.5%; Score 767.5; DB 10; Length 180;  
 Best Local Similarity 75.4%; Pred. No. 7e-45;  
 Matches 144; Conservative 19; Mismatches 17; Indels 11; Gaps 1;

OY 20 IEHSVDLTVEMAMDSQSLDKYERPEVLSMLPRYEGSKYLELAGIGRTGLAEKAS 79  
 DB 1 MEHSDDLTVEMAMDSKASDLKEREPEVLSLIPYEGSKYLELAGIGRTGLAEKAG 60  
 OY 80 QVTAIDFIESYIKNESINGHYKVKFMCADVTSPSLNISNSVDITFSNMLMYSDEE 139  
 DB 61 EVIADLTIESAIQKNEVNGHYKVKFMCADVTSPDLKIDGSLDITFSNMLMYSDEE 120  
 OY 140 VERLVERMLKWLKPGYIFFRESGFHSGDHKRSNPTHYREPRYTKIFKCHMODSG 199  
 DB 121 VELMAERHMGVKGPGYIFFRESGFHSGDHKRSNPTHYREPRYTKV----- 169

OY 200 NSYELSLGCK 210  
 DB 170 QSPDLTPEFLK 180

RESULT 6  
 022993 PRELIMINARY; PRT; 437 AA.

AC Q22993; 01-NOV-1996 (TREMURREL. 01, Created)  
 DT 01-NOV-1996 (TREMURREL. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMURREL. 14, Last annotation update)  
 DE SIMILAR TO METHYLTRANSFERRASES.  
 GN P54D11.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodermidae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson R., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Pauley A., Gattung S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U64834; AAB04824.1; -  
 DR InterPro: IPR000051; -  
 DR InterPro: IPR000339; -  
 DR InterPro: IPR001601; -  
 DR Pfam: PF01209; Ublc\_methyltran; 1.  
 KW Transferrase; Methyltransferase.  
 SQ SEQUENCE 437 AA; 49769 MW; 6675E262F627D08B CRC64;

Query Match 19.2%; Score 498.5; DB 5; Length 437;  
 Best Local Similarity 32.6%; Pred. No. 4.6e-26;  
 Matches 127; Conservative 70; Mismatches 156; Indels 37; Gaps 10;

OY 121 NSVYIFSNMLMYSDEEVLVERMLKMLK-----GGYIFFRESGFHSGDHK 173  
 DB 64 NADVLIFSNMLSOITNADL-----LTDPLKNTATNATIGYIITED-LKDSDRQV 116  
 OY 174 SNPHYREPRYTKIFKCHMODSGNSYELSLGCKCI--GAYVSKKNOQISLWM-- 229  
 DB 117 ARLDY-----FDVFRITDSDGNNTGLDLYTVQVESHNYE---QNFIDFLFV 163  
 OY 230 -QKY---DSEDDKGFORFLDSSQYKFNLSILRYEVPFGYVSTGLTKEFVSKL-DLK 284  
 DB 164 RKVFAPPTDATTFRFLDQTYNTGIDAYEMFGVNFSPGYDENLKIIRFGDFK 223  
 OY 285 PGQKVLVCGIGGSGDYEMAEYDVEVYIDLSINMISFALERSIGLK-CAVEFEVADCT 343  
 DB 224 PGQTMIDIGVIGGAGQVADDEFVHVGIDLSNMALALERHREKDSVKKSIIDAL 283  
 OY 344 KKDYPENSFDVYISRDITLHIDQKPLFRSFHKLKPGKVLISDYCKSACTPAEERAY 403  
 DB 284 VYQFEDNSFDVFSRDCIOHIDPEKLFISRYKALKPGKVLIMYKGGVGEQSDKFKTY 343  
 OY 404 IROGVDLHDVKAQKMLKQAGFEVIAENRTDFOIYLOKELALOEKXDFIDDSSE 463  
 DB 344 VAQRVFLKMLKETADIANKTGFVNVOTENMTPFKETLEERGHLDNAEFSKFTOR 403  
 OY 464 DYNDIVDGWAKALVTEGEOQGLFIATK 493  
 DB 404 ERDSLIGWMDKGLYIEKDNHNMWFLAOK 433

RESULT 7

023552 PRELIMINARY; PRT; 495 AA.

AC Q23552; 01-NOV-1996 (TREMURREL. 01, Created)  
 DT 01-NOV-1996 (TREMURREL. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMURREL. 14, Last annotation update)  
 DE CODED FOR BY C. ELEGANS CDNA YK92B11.3.  
 GN ZK622.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodermidae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson R., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;

Query Match	17.3%	Score 449.5	DB 5	Length 495
Best Local Similarity	28.6%	Pred. No. 1.2e-22		
Matches 137	Conservative 74	Mismatches 195	Indels 73	Gaps 17
QY	12	REVFKKWIEHSVDLTVEBAMLMDSQASDLDRKREPEVLSMLPEPYBGSVLEAGIGRETT	71	
Db	41	RANFESFDKXSDKDPDTSMMLNHSAELESSDRADILASLPDLHNKDVYDAGIGRPT	100	
QY	72	GELAKAQAQVLAADPIESVYKKNESINNGHYKVKVCMCADVYSPSLNISPNVDILFSSNML	131	
Db	101	TVLAETAAWVASTDPIEDSFICKKQERNNAHLGININQVGAV--GLKMSNSVDLVFTNML	158	
QY	132	LMVLSDEEVEERLVERMLMKPGYIFRFESQCFHOS-GDHKKRS-----NPTHYREPR	183	
Db	159	MMVLSDEEVEERLVERMLMKPGYIFRFESQCFHOS-GDHKKRS-----NPTHYREPR	218	
QY	184	FYTKFFECCHMDDSGNSYELSLGCKICIGAYVSKKKNQONISMLMOKYDSEDD-KG--F	240	
Db	219	LYNLLRAIRYRDVYNKLMRFVNWQSCSPPTYIKKSNMNRQVHMLAEKVPADGAKGTSE	278	
QY	241	QRELD-----SSQYKKNISILRYRERVGPGYVSGLETTFFYSKIDLKQGO-----	287	
Db	279	NEIVELINQNMQNEBANDAKLDDK-----YV-----WTDKVFSSALLSLPSNGTFELX	328	
QY	288	--KVLVDGCGGIGGDFYMAENTDVEYVGDLSINMISFALERSIGLK-CAVEF-----	337	
Db	329	TPRFTVSPYCHINA--HTLAETFNANVMWTEILPEYRTSTLTKSNMLKQDRAFGWNSQLT	386	
QY	338	-EVADCTKKDVPENSEFDVYSRDTLTHLODKPALRFRSPFKMYLPGGKVLISYCKSAGTP	396	
Db	387	DSVYTWQOKD---ALFDVFAATEFISTYDDE--TRQLRPVMSDGAKKFTLEPEVBN--	439	
QY	397	SAEFAAYTRQRCQYDLHDVAKYGMKLDGAEVEVIAENRTDQFIQVYLQKELDALEQEKD	455	
Db	440	EAEMKORIOELGYTL-----KSFETDV-----TDQCTEAEQYFQKQHEQLRDE	481	
RESULT	8			
09KUJ20	PRELIMINARY	PRT:	565 AA.	
AC	09KUJ20:			
DF	01-OCT-2000 (Tremblrel. 15, Created)			
DF	01-OCT-2000 (Tremblrel. 15, Last sequence update)			
DF	01-MAR-2001 (Tremblrel. 16, Last annotation update)			
DE	GLYCINE-SARCOSE-DIMETHYLGLYCINE METHYLTRANSFERASE.			
OS	Actinopolyspora halophila.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC	Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;			
OC	Actinopolyspora.			
OX	NCBI_TaxID:1850;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE-20357342; PubMed-10896953;			
RA	Nyysola A., Kervono J., Kaukinen P., von Weymar N., Reinikainen T.;			
RT	"Extreme halophiles synthesize betaine from glycine by methylation.";			
RL	J. Biol. Chem. 275:22196-22201(2000).			
DR	EMBL: AF216283; AAF87204.1; -			
DR	InterPro: IPR000051; -			
DR	InterPro: IPR001601; -			
KW	transferase; Methyltransferase.			

[illegible]

```

OY 280 KLDLPGKGVLDYGGGJGGGFGFYMAENYDVGIDISIMNIS--FALRESGIKACVEF 3377
Db 114 QLELPGKVLVDYGGGJGGJGPREIARSRSSVTSYTLNNNDQYITGKALNNSVGLATCDF 173
OY 338 EVADCTKKDYDENSEFDVYYSRDTLLHIODKPALERFSEFKMLPGKVLISDYC-KSAGTP 3966
Db 174 KKADPMKMPFSDNFPDVAYLAEATCHADPVGCKEILRYLAKPGQCAVYEWCTIDHYDP 2333
OY 397 SAEFAAYLR---QRCYDLADYKAYGKML--KDAQFEVYI---AENR----- 4344
Db 234 NNATKRKDEIEGLGNGLPDIRSTROCLQAVKADGAF-EVIDKDLADESPLPWYLLPDPS 2922
OY 435 -----TDQFQVYQKELDALEQEKDFIDDFSEEDYNDYDGGKAKL 476
Db 293 RFLSSFRLLTVGRIITRNMKVL--EYVGLAPGSGQSVRSFLEKAEGLVEGKKEI 348

RESULT 10
Q41587
AC Q41587 PRELIMINARY; PRT; 363 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 11, Last annotation update)
DE DELTA-24-STEROL METHYLTTRANSFERASE.
GN TA-MT.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticaceae;
OC Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Subramaniam K., Deng P.P.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U60755; AAB49338.1; -.
DR Mendel; 10874; Triae; 1138; 10874.
DR InterPro; IPR000051; -.
KW Transferase; Methyltransferase.
SO SEQUENCE 363 AA; 41636 MW; 7F209958661A62C75 CRC64;

```

Query Match	Similarity	8.0%	Score 207	DB 10	Length 363
Best Local	Similarity	28.6%	Pred. No. 2.2e-06		
Matches	68	Conservative	38	Mismatches	86
				Indels	44
				Gaps	8

  

ID	Q55809	PRELIMINARY	PRT	317 AA
AC	Q55809			
DT	01-NOV-1996	(TREMblrel. 01, Created)		
DT	01-NOV-1996	(TREMblrel. 01, Last sequence update)		
DT	01-MAR-2001	(TREMblrel. 16, Last annotation update)		
DE	DELTA(24)-STEROL C-METHYLTRANSFERASE.			
ES	ERG6 OR SLR0089.			
OS	Synechocystis sp. (strain PCC 6803).			

  

QY	280	KIDLKPGKVLVDVGGCGGGGDFYMAENYDVEVVGIDLSINMIS--FATERSIGLKCAVEF	337
Db <td>114</td> <td>QELKPGKVLVDVGGCGGGGDFYMAENYDVEVVGIDLSINMIS--FATERSIGLKCAVEF</td> <td>173</td>	114	QELKPGKVLVDVGGCGGGGDFYMAENYDVEVVGIDLSINMIS--FATERSIGLKCAVEF	173
QY <td>338</td> <td>EVADCTKKDIPNSNDVIYSRDTIIHIDKPLFESFHKMLKPGKVLISDYC-KSAGIP</td> <td>396</td>	338	EVADCTKKDIPNSNDVIYSRDTIIHIDKPLFESFHKMLKPGKVLISDYC-KSAGIP	396
Db <td>174</td> <td>VKADMKMFSDNTEDAVAYALEATCHADDPVGCYEIKRYLKPGGCAFAYEWCTIDHYDP</td> <td>233</td>	174	VKADMKMFSDNTEDAVAYALEATCHADDPVGCYEIKRYLKPGGCAFAYEWCTIDHYDP	233
QY <td>397</td> <td>SAEFAAYLR--QRGYDLDHVKAYKML---KDAQFEVYI---AEN-----</td> <td>434</td>	397	SAEFAAYLR--QRGYDLDHVKAYKML---KDAQFEVYI---AEN-----	434
Db <td>234</td> <td>NNATKKRIKDELELNGLPDIRSTGQCQAAVKDAEF-EYIMKDLALSDPLFWYLPDPS</td> <td>292</td>	234	NNATKKRIKDELELNGLPDIRSTGQCQAAVKDAEF-EYIMKDLALSDPLFWYLPDPS	292
QY <td>435</td> <td>-----TDQFIQVLOKELDALFEQEKDFIDFSESDYNDYDGMKAFL</td> <td>476</td>	435	-----TDQFIQVLOKELDALFEQEKDFIDFSESDYNDYDGMKAFL	476
Db <td>293</td> <td>RFSLSFRITTVGRITTENMVKVL--EYVGIAPEGCSQRVSSFLERKAAGLVEGKRKEI</td> <td>348</td>	293	RFSLSFRITTVGRITTENMVKVL--EYVGIAPEGCSQRVSSFLERKAAGLVEGKRKEI	348

CC Bacteria|Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96127529; Pubmed=8590279;  
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Suglura M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
RT region from map positions 64k to 92k of the genome.";  
RL DNA Res. 2:153-166(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; Pubmed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Suglura M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nairu K., Okumura S.,  
RA Shimpu S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
EMBL: D64004; BAI10562.1; -.  
DR InterPro: IPR000051; -.  
DR InterPro: IPR001601; -.  
DR Transferrase: Methyltransferase.  
SO SEQUENCE 317 AA; 35571 MW; 98DD01C081B89F72 CRC64;

```

Query Match=          7.9% ; Score 204.5; DB 2; Length 317;
 * Best Local Similarity    26.3%; Pred. No. 2.7e-06;
Matches   61; Conservative 48; Mismatches 90; Indels 33; Gaps

QY  215 YKSKKNONQISWLMQKVSEDDKGFORFLDSSOQKYKNSILRYERYVPGPVSTGLET 274
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   35 YEKIKNFYEDSSGLMEDVMGE-----HMHGYGGPHGTGRIDR-----ROAQIDL 80

QY  275 KEF-----YSKDLKPQGKVLDPVCGCTGGDFPYMAENTDEVGVGDLSINMISTALE--RS 328
      ||| :|| :|| :||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   81 KELLAMAVPQSNAP-RRIIDLDGGCIGGSLLYLAAQQQAPEVWGASLSPVERAGERAR 139

QY  329 IGLCAVEFEVDATCKRKDPENSEDVYTSRDITLIHQDKALFRSFKMKLPGBKVLYSD 368
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   140 LGLSGTCGFVANMLDLRFADSFSDWMWSLESNGHMNKAGQLQEAARYLVAKPSGRILLAI 199

QY  389 YC----KSACTPSAEPFAAIVIRQDYDL-----DYKAYGKMICKDAEYEV 429
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   200 WCHRPIDPGNPLTADERRH-L-QAIYDVYCLIPVYSLPDYEATAIARECGEGEI 250


RESULT  12
QREV12 PRELIMINARY; PRT; 278 AA.
AC QREV12;
DT 01-MAR-2001 (TREMBREL_16, Created)
DT 01-MAR-2001 (TREMBREL_16, Last sequence update)
DT 01-MAR-2001 (TREMBREL_16, Last annotation update)
DE SNOGM.
GN SNOGM.
OS Streptomyces nogalater.
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=38314;
[1]
RN SEQUENCE FROM N.A.
RA STRAIN=ATCC27451;
RA Torkkell S., Kuninari T., Palmu K., Manttala P., Hakola J.,
RA Ylthonko K.;
RT "The entire nogalamycin biosynthetic gene cluster of Streptomyces
RT nogalater: Characterization of a 20 kb DNA region and generation of
RT hybrid structures."
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
```

DT 01-NOV-1996 (TREMBLref. 01, Created)  
DT 01-NOV-1996 (TREMBLref. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBLref. 14, Last annotation update)  
DE S-ADENOSYL-L-METHIONINE:DELTA24-STEROI-C- METHYLTRANSFERASE.  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eurosids I;  
OC Fabales; Fabaceae; Papilionoideae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP SPRAIN-CV. WILLIAMS 82; TISSUE-ETIOLATED HYPOCOTYLS;  
RC MEDLINE=96199190; PubMed=8621604;  
RX Shi J., Gonzales R.A., Bhattacharyya M.K.;  
RA "Identification and characterization of an S-adenosyl-L-methionine  
RT delta 24-sterol-C-methyltransferase cDNA from soybean.";  
RL J. Biol. Chem. 271:9384-9389(1996).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP SPRAIN-CV. WILLIAMS 82; TISSUE-ETIOLATED HYPOCOTYLS;  
RC Clouse J.A.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

